TSUNAMI HAZARD ASSESSMENTS WITH CONSIDERATION OF UNCERTAIN INPUTS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Ignacio Sepulveda Oyarzun

December 2017

TSUNAMI HAZARD ASSESSMENTS WITH CONSIDERATION OF UNCERTAIN INPUTS

Ignacio Sepulveda Oyarzun, Ph.D.

Cornell University 2017

Tsunami hazard assessments have frequently been conducted by means of numerical models using deterministic inputs, such as the earthquake fault parameters and bathymetry. Some fault parameters of the earthquake, however, can only be estimated probabilistically for a future event. The bathymetry, on the other hand, is surveyed with an imperfect resolution and accuracy, which can lead to errors in the tsunami modeling results. Hence, tsunami assessments before an event are unavoidably uncertain. This thesis aims to describe how the uncertainties of some earthquake fault parameters and bathymetry impact the tsunami assessment uncertainties.

The earthquake uncertainties are related with the assumed aleatory nature of the slip distribution and the rupture location, while the bathymetry uncertainties are associated with the lack of data in unsurveyed areas. These uncertain inputs and the corresponding tsunami response are modeled as random elements by adopting a stochastic approach. The uncertain earthquakes of this thesis consider a slip distribution modeled as an homogeneous random field and a rupture location modeled as a random vector. The bathymetry, on the other hand, is modeled as a non-homogeneous Gaussian random field, which is conditional to surveyed data. The generation of samples of the earthquake rupture location is straightforward. Conversely, the generation of samples of the slip distribution and bathymetry are rather complex. By means of a Karhunen-Loeve expan-

sion and a translation model we propose a consistent method for the generation of samples of these uncertain inputs. Unlike other approaches, the Karhunen Loeve expansion generates consistent and accurate samples of non-rectangular random fields. The uncertainties of tsunami hazard assessments are then quantified by means of a Stochastic Reduced Order Model (SROM), which is more accurate than the classic Monte Carlo simulation for a same number of samples.

The uncertainty quantification methods developed in this thesis are presented with two illustration cases. In one illustration case we study M_W 8.0 earthquakes within a seismogenic region in North Chile. First, we demonstrate that our proposed method generates consistent earthquake and bathymetry samples. Second, we demonstrate that estimates of tsunami assessment uncertainties obtained with SROM are more accurate than estimates obtained with classic Monte Carlo simulations. From sensitivity analyses and comparison with records of the 2014 earthquake tsunami, we also conclude that the probability properties of the analyzed uncertain inputs and other aspects of the tsunami assessment can be relevant sources of uncertainty. In the second illustration case we perform a probabilistic tsunami hazard assessment (PTHA) which assesses earthquakes generated in the Manila Subduction Zone and tsunami responses in Hong Kong, China, and Kao Hsiung, Taiwan. First we demonstrate that our proposed methods can be combined with PTHA to account for the earthquake fault parameters and bathymetry uncertainties. Second, we demonstrate that the earthquakes recurrence model and the tsunami propagation model adopted in the PTHA constitute additional sources of uncertainty, which can be as relevant as the uncertainties of the earthquake fault parameters and bathymetry.

BIOGRAPHICAL SKETCH

Ignacio Sepulveda Oyarzun born in 1984 in Valdivia, Chile. He earned his degree of Civil Engineer from Santa Maria University (Chile) in 2011. After his graduation, he worked as a lecturer and researcher of hydrodynamics and numerical modeling in the School of Ocean Engineering at Valparaiso University. He also worked as a consultant engineer in AWAS Ingenieria. In August 2013 he joined the MSc/PhD program of the School of Civil and Environmental Engineering at Cornell University, with funding support from Fulbright Commission and Becas Chile. While studying at Cornell, Ignacio worked on many topics regards to tsunamis, environmental fluid mechanics, stochastic calculus and seismology. His main research regards to the uncertainty quantification in tsunami hazard assessments. Some of his research has been published in the Journal of Geophysical Research Letters, the Journal of Coastal Engineering and the Journal of Geophysical Research-Solid Earth. After graduation, Ignacio plans to join a post-doctoral program on extreme coastal events.

This document is dedicated to Alejandra.

ACKNOWLEDGEMENTS

The application cases of this thesis used topo-bathymetry data provided by the Marine Dept. and the Civil Engineering and Development Dept. of Hong Kong, Dr. Tso-Ren Wu from Central University of Taiwan and the Chilean Navy.

I would like to thank Prof. Philip Liu. His experience and judgment taught me valuable lessons which forged me as a scientist and researcher. I will never forget two important advises he gave me during these four years: be a curious person and you must criticize your own research as the toughest reviewer. Thanks to be my mentor. I would like to thank Prof. Mircea Grigoriu, who invited me to explore stochastic approaches for uncertainty quantification. He always knew the simplest way to explain things. It was an honor to be a member of his research group. I would like to thank Prof. Matthew Pritchard. It is difficult to find a person involved in so many topics like him. He always provided excellent suggestions and comments to improve my research. I think I was very fortunate to have this "dream team" as committee. I would like to thank all EFMH faculties, especially Prof. Diamessis and Prof. Cowen. I had the lucky to take their courses and learn more about turbulence, stratified flows and experimental methods. Thanks to them, I could improve my understanding of environmental fluid mechanics.

I would also like to thank all the friends I made in CEE and Ithaca. I especially would like to thank Gustavo, Pato Winckler, Blair, Jeff, Diego, Mahmoud and all EFMH's. They were an important support for me and my family. We never felt alone. I also would like to thank all the support of the staff of CEE.

I want to sincerely thank to my wife, Alejandra, and my son Agustin. They generously share their time with my desire to be a Doctor. Finally, I would like to thank my parents, siblings, uncles, aunts and cousins for all their support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

A relevant topic in geophysics and oceanography is the study of tsunamis. Tsunamis are trains of long waves which propagate in large water bodies such as oceans and lakes. They are generated by large perturbations, such as earthquakes and landslides. Due to shoaling processes, these waves generate inundation at coastal areas, which is frequently associated with human life and property losses. The tsunamis studied in this thesis are generated by the deformation of the seafloor, which is caused by the dislocation within a fault during an earthquake. The assessment of these tsunamis relies on models of the physical phenomena involved and observations used as input data.

Tsunami models are composed of a tsunami generation model and a tsunami propagation model. It is well-known that during an earthquake the strain between two contacting tectonic plates, which have been accumulated during an inter-seismic period, is suddenly released, and the rupture occurs. The rupture area (denoted by R in the left panel of Fig. 2.1) is located within a seismogenic region. The earthquake induced ground motion dynamics is complex. However, it has been demonstrated that seabed deformation can be reasonably described by the Okada model (Okada, 1985), which depends on the displacement between the tectonic plates and the geometry of the rupture area. Adopting further assumptions, the seafloor movement can be viewed as an impulsive motion and the vertical seafloor displacement pushes the entire water column above upwards instantaneously. Consequently, the shape of the initial water surface deformation mimics that of the vertical seafloor deformation (Todorovska and Trifunac, 2001). Once the free water surface is deformed, this disrupts the hy-

drostatic equilibrium in the water body and hence, tsunami waves are generated due to gravity. Further details about the assumptions and simplifications of the tsunami generation model are provided in Chapter 2. The wavelength of large tsunami leading waves is $O(100km)$, which is very long compared with the averaged ocean depth, $O(1-3km)$ (An and Liu, 2014). Furthermore, the amplitude of leading tsunami waves in the ocean basin is $O(1m)$ and is very small in comparison with water depths. Therefore, the nonlinearity can be neglected and the linear shallow water wave equations are adequate in modeling the leading tsunami waves propagation in the ocean basin (e.g. Mei et al. (1989)). However, as leading tsunami waves propagate onto continental shelves and coastal areas, wave amplitudes grow significantly because of wave shoaling. Hence, the non linear inertia terms should be included in the shallow water wave equations model so as to simulate the nonlinear tsunami behaviors in shallow regions. It is important to remark that the tsunami propagation is highly controlled by the bathymetry and, therefore, it is a relevant input in the non linear shallow water wave equations. There are several existing tsunami propagation models available, such as TUNAMI-N2 (Shuto et al., 2006), MOST (Titov and Gonzalez, 1997), COMCOT (Wang, 2009) and NEOWAVE (Yamazaki et al., 2012). Once tsunamis are simulated in the mentioned models, information such as surface elevation, velocities and runup can be determined. Further details of the shallow water equations are provided in Chapter 2.

The tsunami hazard assessments are inherently uncertain due to the existence of various sources of uncertainty in the tsunami modeling process. Fig. 1.1 shows a chart with sources of uncertainty which are commonly relevant in tsunami hazard assessments. Input data (box A in Fig. 1.1) is uncertain due to errors in the measurement of some data and the aleatory nature of others. The

models themselves (box B in Fig. 1.1) also contribute with uncertainty because of the assumptions and simplifications adopted on them. These latter are further described in Chapter 2. In addition to the sources of uncertainty in tsunami modeling, the interaction of tsunami waves with other coastal processes (box C in Fig. 1.1) can also contribute with uncertainty. Tides, for instance, are long waves associated with temporal changes of sea level and currents, which might significantly influence tsunami responses. Since the tidal phase during a future tsunami event is uncertain, tsunami responses may be uncertain as well. This thesis addresses the uncertainty of two relevant inputs in tsunami assessments. They are the aleatory uncertainty of the characteristics of future earthquakes and the epistemic uncertainty of bathymetry when available data is insufficient. The effects of these uncertainties on tsunami hazard assessments are then quantified.

In the last decades it has arisen an interest to assess the potential hazard of tsunamis, especially close to subduction zones. There are two typical approaches for assessing tsunami inundation hazard. The first approach is the so called "worst case scenario assessment" in which tsunami responses to an estimated maximum earthquake magnitude with a set of fault parameters are analyzed (e.g. González et al. (2005)). This approach is deterministic and does not address the uncertainty of the input data. The second common approach is the so called "probabilistic tsunami hazard assessment" (PTHA) (e.g. Geist and Parsons (2006)) in which the likelihood of exceeding some level of tsunami responses (such as tsunami wave amplitude or runup height) is determined over a specified period of time (i.e., 100 years or 500 years). The PTHA approach adopts a probabilistic model for the earthquake recurrence in terms of magnitude. The uncertainty of input data can be further considered in the PTHA

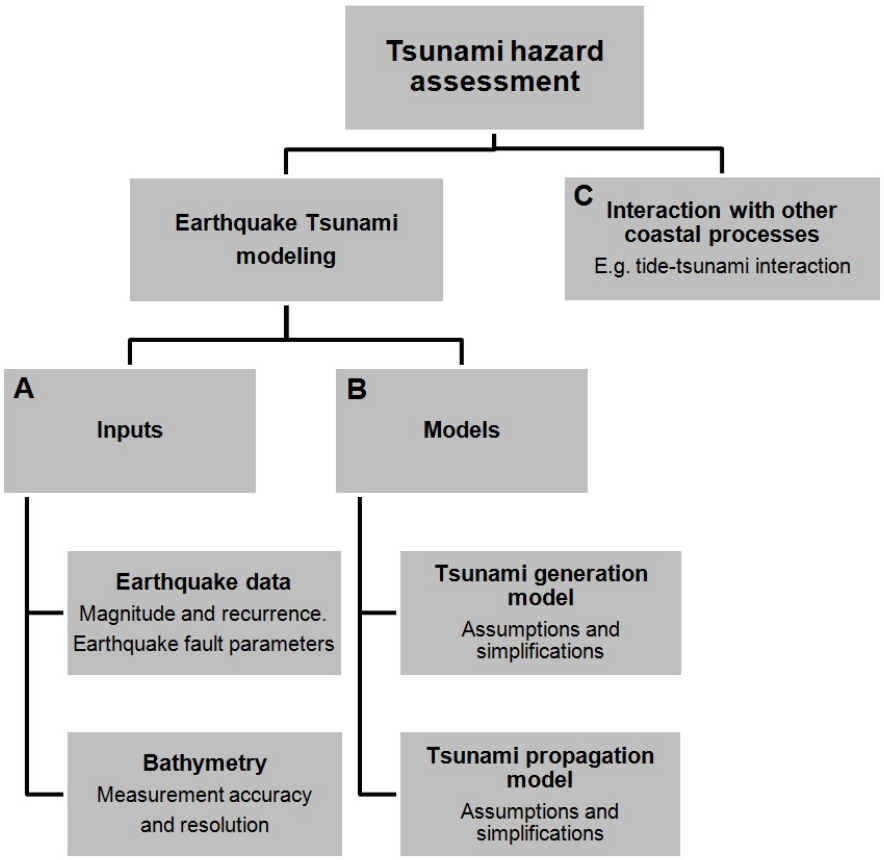


Figure 1.1: Relevant sources of uncertainty in tsunami hazard assessments.

approach. However, its inclusion is often incomplete, or simply omitted, due to difficulties in describing consistent properties of the input uncertainties, and relating them with tsunami response uncertainties. Hence, these commonly used hazard assessment approaches may lead to an inaccurate perception of risk.

To predict and characterize the fault mechanism of a future earthquake is extremely challenging with high uncertainty. With exception to some rare earthquake mechanisms reported in the literature (e.g. 1992 Nicaragua earthquake (Satake, 1994), 2007 Kuril Islands earthquake (Ammon et al., 2008) and 2009 Samoa-Tonga earthquake (Lay et al., 2010)), some fault parameters can be con-

strained or specified with acceptable accuracy in subduction zones, such as those related to the plane geometry. However, two important fault parameters, slip distribution of the earthquake and fault location, are rather unpredictable. By adopting an stochastic approach, these fault parameters can be modeled as random elements, with consideration of scaling laws relating probability properties with earthquake magnitude.

Another relevant input for tsunami hazard assessments is the bathymetry data, which controls the tsunami propagation. Tsunami propagation models require bathymetry data in specific locations, known as grid (mesh) nodes. The grid resolution is defined for each particular tsunami model so as to capture relevant length scales of the tsunami wave during its propagation (e.g. the wavelength of tsunami waves). The bathymetry data, on the other hand, is acquired in specific point locations and with a specified spatial density, which rarely coincides with the location and density of grid nodes. Hence, interpolation methods has to be adopted to estimate the bathymetry of the grid nodes. The bathymetry at the grid nodes has two relevant sources of uncertainty: errors of measured data and interpolation errors. In this thesis we address interpolation errors since this uncertainty source is expected to be dominant. As the case of the uncertainties of earthquake characteristics, we adopt a stochastic approach to model bathymetry uncertainties as a random element, with consideration of surveyed data.

The final goal of this study is to quantify uncertainties in tsunami hazard assessments. Therefore an uncertainty propagation model is implemented in the framework of this thesis, which uses the modeled input data uncertainties. The Stochastic Reduced Order Model (SROM) (Grigoriu, 2009) is used to propagate

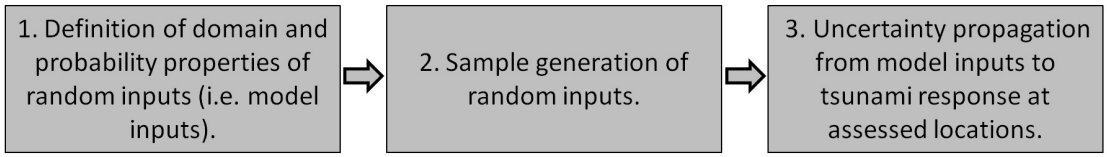


Figure 1.2: Flow chart of stochastic approaches that model uncertain inputs and propagate uncertainty into tsunami hazard assessments.

the uncertainties, which is further described later. The sequence of the stochastic approach assessing the uncertainty of tsunami assessments due to uncertain input data is resumed in the flow chart of Fig. 1.2. Each stage faces significant challenges, which will be discussed in Chapter 2. The objectives of this thesis are to develop solutions to these challenges and offer a consistent framework to quantify uncertainties in tsunami hazard assessments.

The methodology proposed in this thesis is applied on real case problems. First, we use an illustration case in Northern Chile with consideration of earthquakes at the same seismogenic region of the 2014 Chilean earthquake. This illustration case is employed to show the methodologies addressing uncertain earthquake characteristics and bathymetry. Finally, in Chapter 5 we assess uncertainties of a probabilistic tsunami hazard assessment in South China Sea with consideration of uncertain earthquakes at the Manila Subduction Zone.

CHAPTER 2

THEORETICAL FRAMEWORK

2.1 Earthquake tsunami generation

In this study we shall focus in earthquakes induced in subduction zones, which are the largest earthquakes recorded and more relevant in the context of tsunamigenic potential. These are interplate earthquakes with a reverse fault mechanism. In their set up, a plate moves and sinks below the other. In the plate interface shear stresses, and thus potential energy, is accumulated during inter-seismic periods. In most of the cases this energy is released in earthquakes. The released energy during an earthquake event is defined as seismic moment, M_0 (Aki, 1966), and is given in terms of the fault parameters,

$$M_0 = \mu_E R \bar{S}_R, \tag{2.1}$$

where R is the rupture area located within a seismogenic region, as shown in the left panel of Fig. 2.1. The seismogenic region is defined as the area of the plate interface that undergoes co-seismic slip during thrust earthquakes (e.g. Scholz and Campos (2012) and Kozdon and Dunham (2013)), and the region may be bounded by the earth surface (so called trench in subduction zone earthquakes) and regions where aseismic deformation occurs. In Eq. 2.1, μ_E is the crust rigidity and takes the value of 40 GPa, according to the proposed values of Kagan (1997) and Bird and Kagan (2004). $\bar{S}_R$ is the spatially averaged slip over the rupture area. It is important to remark here that the fault slip is spatially varying (heterogeneous) and is often called the slip distribution. Earthquakes are also measured by means of their magnitudes. In particular, the Moment Magnitude M_W is related with M_0 by (Kanamori, 1977),

$$M_W = \frac{2}{3}\left[log_{10}(M_0) - 9.1\right], \tag{2.2}$$

with M_0 in Nm units. The seismic moment and moment magnitude quantify the total energy released on a earthquake but do not provide the detail of the rupture area and slip magnitudes. The stress drop, defined as $\Delta\sigma$, quantifies the shear stress released in a seismic event. It is given by the ratio of slip and rupture area,

$$\Delta\sigma = C\mu_E \frac{\bar{S}_A}{\sqrt{A}}, \tag{2.3}$$

where C is a constant, A is any sub-region in the rupture area and $\bar{S}_A$ is the mean slip in that sub-region. The stress drop typically varies between 1 and 100 MPa (Allmann and Shearer, 2009), which indicates that the maximum slip magnitude is bounded.

The dimensions of the rupture area follow scaling relations with the earthquake magnitude (e.g. Kanamori and Anderson (1975); Wells and Coppersmith (1994); Mai and Beroza (2000); Blaser et al. (2010); Strasser et al. (2010)). The rupture plane dimensions L and W, and the area of R are related to the moment magnitude M_W in the following form, $log(X) = a + bM_W$, in which X is L, W or area R, respectively, and a and b are empirical constants. The scaling relations developed by Blaser et al. (2010) and Strasser et al. (2010) are particularly suitable for the purposes of our study since they, independently, analyzed tsunamigenic thrust earthquakes in different seismogenic regions. A comparison of the scaling relations in Fig. 2.2 shows small differences, especially for $M_W > 8.0$.

As we mentioned in the introduction, two important fault parameters, slip distribution of the earthquake and rupture location, are rather undetermined for future earthquakes. The slip is heterogeneous and depends on the hetero-

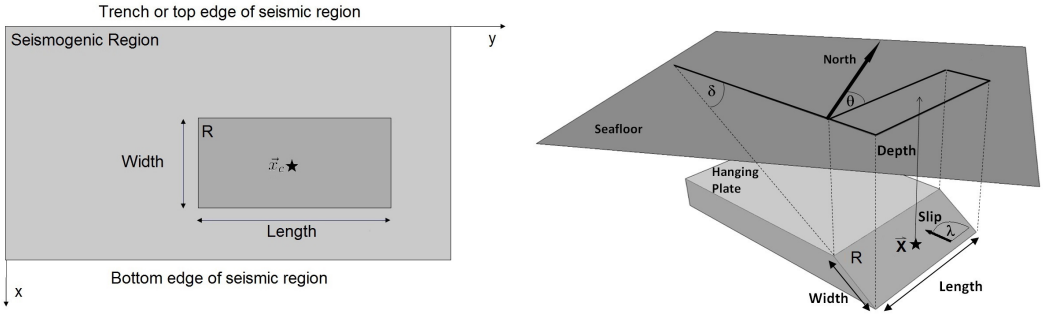


Figure 2.1: Idealized fault geometry. Left panel: Seismogenic region and rupture area R in the interface of two tectonic plates. The coordinate x and y point in the dip and strike direction, respectively. The star indicates the earthquake centroid location $\vec{x}_c$. Right panel: Relevant fault parameters for the determination of the seafloor displacements. Fault plane width W, Fault plane length L, Fault depth d, Slip variable over the rupture area S, strike angle θ, dip angle δ and rake angle λ.

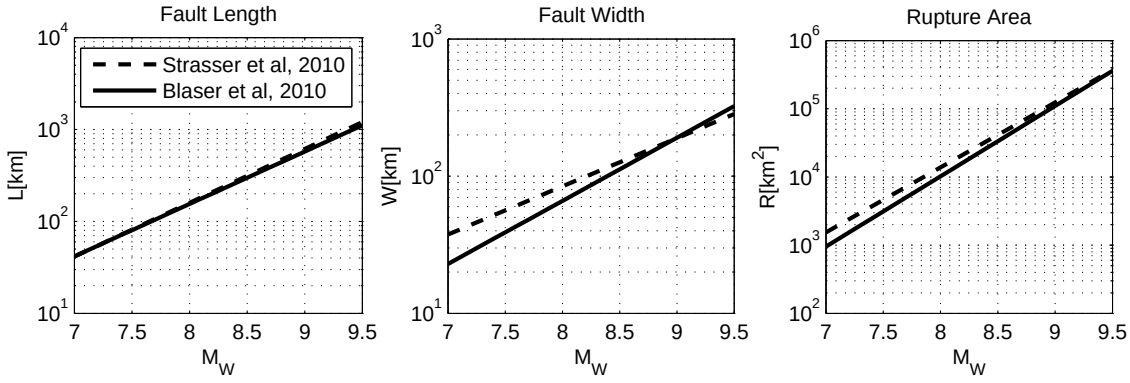


Figure 2.2: Scaling relation curves for L, W and rupture area R of thrust faults. The rupture area of Blaser et al. (2010) was obtained by the product of L and W.

geneities of plate interface stress state, rock properties and fault geometry [Andrews, 1981]. The rupture location within a seismogenic region, on the other hand, has been suggested to depend on the heterogeneity of rock properties, the normal stresses on the fault and temperature [Mori and Abercrombie, 1997]. The complex interactions between the aforementioned factors, make the prediction of slip distribution and rupture location for future events unfeasible. This thesis addresses the modeling of these uncertain fault parameters by adopting a stochastic approach.

It is noteworthy to mention that rare small earthquakes generating considerable tsunami inundation have been reported in the literature. The so-called "tsunami earthquakes" are shallow earthquakes which generate anomalously large tsunamis. Examples of this type of events are the 1992 Nicaragua earthquake (Satake, 1994) and the 1896 Sanriku earthquake (Tanioka and Seno, 2001). Different hypotheses to explain such a large tsunamis have been provided in the literature. For instance, Kanamori (1972) suggested that shallow rupture areas might be located within a region of accumulated sediments above the trench. This region is known as the accretionary wedge. It has been observed that rupture speeds are slower and slips are larger in this region, as compared with deeper earthquakes of the same magnitude. As a consequence, the seafloor deformation and the generated tsunami are bigger for a given earthquake magnitude. Tanioka and Seno (2001) also suggested that sediments in the accretionary wedge might experiment more deformation as compared with that of a crust with linear and homogeneous mechanical properties, resulting in a larger generated tsunami. Another type of rare events are the intraplate earthquakes occurring above the subduction interface. These earthquakes are shallower and may contribute with larger seafloor deformations and tsunami wave heights

(e.g. Ammon et al. (2008); Lay et al. (2010)). The prediction of the mentioned rare events is very complex because the seismogenic regions, fault geometries and earthquake magnitudes are not clear or well studied. These rare events are not studied in this thesis.

The tsunamis analyzed in this thesis are generated by the co-seismic seafloor displacement during earthquakes. This displacement is originated by the crustal strain change around the rupture area. As we mentioned in the introduction, it has been demonstrated in the literature that seafloor displacements can be reasonably estimated by the Okada model, which is a linear elastic model of a finite fault in a homogeneous and isotropic half space (Okada, 1985). The Okada model provides analytical solutions in terms of the set of fault plane parameters: length (L), and width (W) of the rupture plane, slip (S), depth (d), dip angle (δ), strike angle (θ) and rake angle (λ) (see right panel of Fig. 2.1). It is important to remark that the assumptions of the Okada model might not hold under some conditions. For instance, rock mechanical properties might be different at the trench because of the accumulation of sediments in the accretionary wedge (Tanioka and Seno, 2001) and, thus, the seafloor deformation of shallow earthquakes might not be accurately modeled by the Okada model. Hence, the validity of the Okada model might be limited in some cases.

The seafloor displacement, generated by an earthquake, perturbs the column of water above it. Consequently, a tsunami is initiated over the earthquake location. The determination of the tsunami initial condition given a prescribed seafloor displacement is complex. However, under the assumptions that (1) water is rather incompressible, (2) the slip and seafloor displacement time scales are much shorter than the time scale of the tsunami propagation, (3) the

characteristic horizontal length of the seafloor deformation is larger than water depths and (4) the horizontal displacement of the seafloor does not contribute in the tsunami generation significantly, the seafloor vertical displacement can be viewed as an impulsive motion, pushing (pulling) the water column upwards (downwards) instantaneously. Consequently, the shape of the water surface initial deformation mimics that of the seafloor vertical displacement. We now give some comments on these assumptions. The assumption (2) is commonly reasonable because rupture speeds are in the order of 2000 m/s, while shallow water wave celerities are in the order of 200 m/s in deep waters. However, rupture speeds in the accretionary wedge might be slower and comparable to shallow wave celerities. The assumption (3) is reasonable for most of large earthquakes. Kajiura (1963) and Geist and Dmowska (1999) suggested that the horizontal length scales of the seafloor displacement should be three times larger than the water depth. We have observed that this threshold might be achieve for the case of shallow earthquakes near the seafloor. To illustrate this, in Fig. 2.3 we compare two solutions of the vertical seafloor deformation, obtained with the Okada model, for a shallow and a deep M_W 8.0 earthquakes with same fault parameters (i.e. same slip, length, width, dip angle, strike angle and rake angle). The seafloor displacement on the left panel corresponds to an earthquake with an up-dip extent to a depth of 1 km and the seafloor displacement on the right panel corresponds to an earthquake with an up-dip extent to a depth of 33 km (i.e. deeper). As we observe in Fig. 2.3, the seafloor displacement corresponding to the shallower earthquake (left panel) has shorter horizontal length scales, as compared with the seafloor displacement of a deeper earthquake with same fault parameters. It is important to mention that the horizontal length scales of the seafloor displacement of the shallower earthquake is about 0.1 degree ($\sim$ 10

km), which is three times the water depth (~ 3 km in this case). The illustration and application cases of this thesis in North Chile and the Manila Subduction Zone consider rupture areas with an up-dip extent to a minimum depth of 1 km and, hence, assumption (3) holds. The assumption (4) holds when the seafloor slope is rather small (Iwasaki, 1982; Tanioka and Satake, 1996). These conditions are often held for deep earthquakes. However, Tanioka and Satake (1996) and Bletery et al. (2015) observed that seafloor slopes close to the trench and horizontal displacement of shallow earthquakes might be significant and contribute to the tsunami generation. They model the contribution in some application problems by assuming an impulsive horizontal movement of the seafloor and by neglecting the injection of horizontal momentum. The modeled contribution was defined as "bathymetry effect" by Bletery et al. (2015). Furthermore, Song et al. (2017) recently suggested that the injection of horizontal momentum to the tsunami initial condition might be relevant as well. Investigations of the tsunami generated by the 2011 Japan earthquake have pointed out that the seafloor horizontal displacement contributed significantly to the tsunami generation, both due to the "bathymetry effect" (Satake et al., 2013) and due to the injection of horizontal momentum (Song et al., 2017). Hence, assumption (4) might not hold for some shallow earthquake cases.

Hereinafter we adopt the Okada model (with homogeneous rock properties) and the four assumptions regarding to the tsunami generation. Hence, the initial water level mimics the vertical seafloor deformation. The depth-averaged horizontal velocities, on the other hand, are initially set to zero.

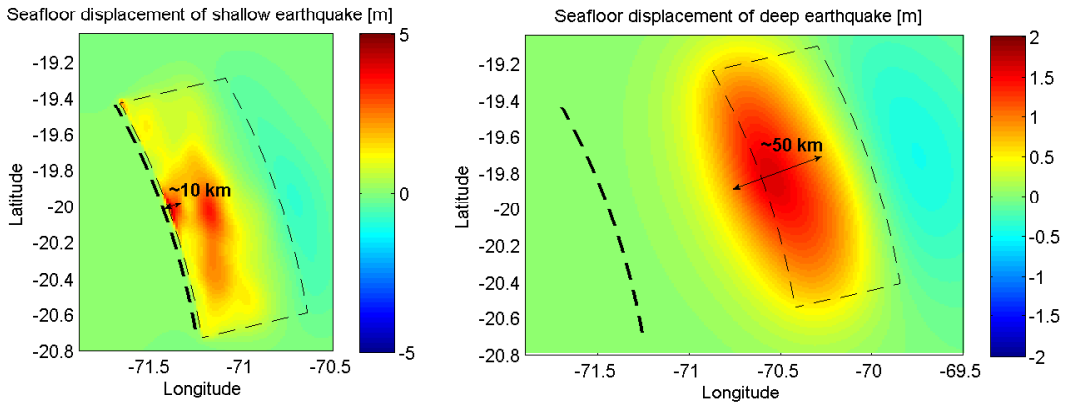


Figure 2.3: Two vertical seafloor displacement solutions for magnitude M_W 8.0 earthquakes. The left panel is associated with a shallow earthquake (1 km minimum depth) and the right panel is associated with a deep earthquake (33 km minimum depth). The thin segmented lines indicates the location of the rupture area, while the thick segmented line indicates the trench. The double sided arrow lines indicate characteristic horizontal lengths of the seafloor displacement.

2.2 Tsunami wave propagation

As pointed out in the introduction, the tsunami wave propagation is modeled by the non linear shallow water wave equations. In this section we first describe the governing equations and boundary conditions. We then describe an analytical solution for a simplified problem and a general numerical model called COMCOT.

2.2.1 Nonlinear shallow water wave equations

Fluid flows can be described by the conservation laws of mass and momentum, which are known as the Navier-Stokes Equations. In this section we present the equations by adopting the Eulerian approach.

The mass conservation law states that a material element following the flow has a constant mass. Thus, the conservation of mass is given by the continuity equation,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\mathbf{v}\rho) = 0, \tag{2.4}$$

where $\mathbf{v}$ is the velocity vector with components (u, v, w) and ρ is the water density. Eq. 2.4 means that the rate of change of mass in the material region is zero. Expanding the equation becomes,

$$\frac{\partial \rho}{\partial t} + \mathbf{v} \cdot \nabla \rho + \rho \nabla \cdot \mathbf{v} = \frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} = 0 \quad \Rightarrow \quad \frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v}. \tag{2.5}$$

As a simplification, it is assumed that the fluid is incompressible inside any material region. That is, $\frac{D\rho}{Dt}$ is zero. With this simplification the continuity equation is obtained for incompressible flows in Eq. 2.6. It is important to note that this equation allows the variation of density in space and time at a given location in the domain, as long as Eq. 2.6 holds.

$$\frac{\partial \rho}{\partial t} = -\mathbf{v} \cdot \nabla \rho \qquad \text{or} \qquad \nabla \cdot \mathbf{v} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \tag{2.6}$$

The momentum conservation law states that the change of linear momentum of a material region is equal to the sum of all the forces acting on it. The forces acting over a material region are body forces, such as gravity, and surface forces, such as pressure and viscous stresses. The momentum conservation equation,

without internal sources of momentum, is given by,

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\right) + 2\Omega\times\mathbf{v} = \rho\mathbf{g} - \nabla P + \nabla\cdot\tau_D, \tag{2.7}$$

where $\mathbf{g}$ is the gravitational force vector per unit of mass (acceleration), P is the pressure, and τ_D is the deviatoric stress tensor (i.e. further stresses at the material region boundaries). The last term in the left hand side is the Coriolis term, which accounts for flows over a sphere and Ω is the rotation angular velocity of the Earth. By assuming a Newtonian fluid and an incompressible flow, $\nabla\cdot\tau_D$ is expressed as,

$$\nabla\cdot\tau_D = \mu\nabla^2\mathbf{v}, \tag{2.8}$$

with μ being the dynamic molecular viscosity. It is important to mention that this viscous stress term is dissipative. By considering Eq. 2.8, Eq. 2.7 adopts the name of Incompressible Navier Stokes Equation.

Now, assume that viscous stresses (and dissipation) are negligible as compared with the Coriolis effect, pressure and gravitational forces. Thus, Eq. 2.7 reduces to,

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\right) + 2\Omega\times\mathbf{v} = \rho\mathbf{g} - \nabla P. \tag{2.9}$$

To solve the conservation laws of mass and momentum in Eq. 2.6 and Eq. 2.9, we consider one dynamic boundary condition and two kinematic boundary conditions,

$$P_{(\eta)} = 0 \quad \text{at } z = \eta \tag{2.10}$$

$$\frac{\partial \eta}{\partial t} + u_{(\eta)}\frac{\partial \eta}{\partial x} + v_{(\eta)}\frac{\partial \eta}{\partial y} = w_{(\eta)} \quad \text{at } z = \eta \tag{2.11}$$

$$\frac{\partial b}{\partial t} + u_{(-b)}\frac{\partial b}{\partial x} + v_{(-b)}\frac{\partial b}{\partial y} = -w_{(-b)} \quad \text{at } z = -b. \tag{2.12}$$

where z is the vertical coordinate, η is the position of the water surface and $-b$ is the position of the seafloor (see Fig. 2.4). We then integrate the conservation of

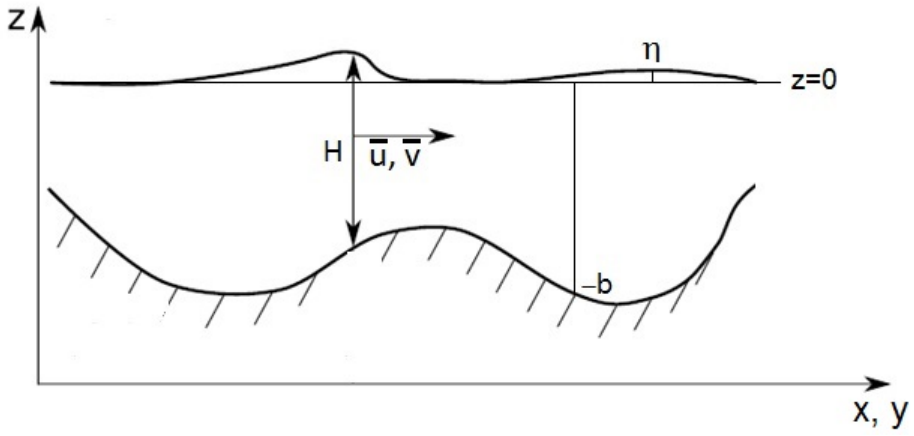


Figure 2.4: Variables of the non linear shallow water wave equations.

mass and momentum equations along z, using η and $-b$ as limits of integration. By defining the depth averaged velocities, $\bar{u}$ and $\bar{v}$, the continuity equation is expressed as,

$$\frac{\partial \eta}{\partial t} + \frac{\partial H\bar{u}}{\partial x} + \frac{\partial H\bar{v}}{\partial y} = -\frac{\partial b}{\partial t}. \tag{2.13}$$

where $H = b + \eta$ is the total water depth. By further considering that $\bar{u} \approx u$ and $\bar{v} \approx v$ for long waves (Mei et al., 2005), the horizontal components of the conservation of momentum equations are expressed as,

$$\begin{aligned} \frac{\partial H\bar{u}}{\partial t} + \frac{\partial H\bar{u}^2}{\partial x} + \frac{\partial H\bar{u}\bar{v}}{\partial y} - fH\bar{v} &= -\frac{1}{\rho}\int_{-b}^{\eta} \frac{\partial P}{\partial x} dz \\ \frac{\partial H\bar{v}}{\partial t} + \frac{\partial H\bar{u}\bar{v}}{\partial x} + \frac{\partial H\bar{v}^2}{\partial y} + fH\bar{u} &= -\frac{1}{\rho}\int_{-b}^{\eta} \frac{\partial P}{\partial y} dz, \end{aligned} \tag{2.14}$$

where $f = 2||\Omega||\sin(\phi)$ is the Coriolis parameter and ϕ is the Earth latitude location.

The vertical component of the conservation of momentum equation takes into account that long waves have small vertical velocities and large wave periods (or characteristic time of the wave). Then, we can assume that the total

derivative of the vertical velocity is negligible, which implies the assumption of hydrostatic pressure. Using the boundary condition of Eq. 2.10 we obtain,

$$\int_{-b}^{\eta} \frac{\partial P}{\partial z} dz = \int_{-b}^{\eta} \rho g dz \Rightarrow P_{(-b)} = \rho g(\eta + b). \tag{2.15}$$

Furthermore, since the total derivative of w is zero along z, we can generalize to,

$$P_{(z)} = \rho g(\eta - z), \quad \text{for} \quad z < \eta. \tag{2.16}$$

Replacing Eq. 2.16 into Eq. 2.14, we obtain the momentum shallow water wave equations,

$$\begin{aligned} \frac{\partial H\bar{u}}{\partial t} + \frac{\partial H\bar{u}^2}{\partial x} + \frac{\partial H\bar{u}\bar{v}}{\partial y} - fH\bar{v} &= -gH\frac{\partial \eta}{\partial x} \\ \frac{\partial H\bar{v}}{\partial t} + \frac{\partial H\bar{u}\bar{v}}{\partial x} + \frac{\partial H\bar{v}^2}{\partial y} + fH\bar{u} &= -gH\frac{\partial \eta}{\partial y}. \end{aligned} \tag{2.17}$$

The variables involved in the equations of above are shown Fig. 2.4. Note that we have assumed that (1) the flow is incompressible, (2) the viscous stress terms (dissipative) are negligible as compared with gravity and pressure terms and (3) the wave is very long, so vertical velocities are negligible and horizontal velocities are constant in depth. It is important to note that the assumption of negligible dissipative stress terms can be further relaxed by adopting simplified models in the momentum equations. For instance, the bottom friction dissipation is the result of shear stresses at the bottom and can be modeled using the Manning quadratic law (Te Chow, 1959). The final form of Eq. 2.17 is given by,

$$\begin{aligned} \frac{\partial H\bar{u}}{\partial t} + \frac{\partial H\bar{u}^2}{\partial x} + \frac{\partial H\bar{u}\bar{v}}{\partial y} - fH\bar{v} &= -gH\frac{\partial \eta}{\partial x} - \frac{gn^2\bar{u}\sqrt{\bar{u}^2 + \bar{v}^2}}{H^{1/3}} \\ \frac{\partial H\bar{v}}{\partial t} + \frac{\partial H\bar{u}\bar{v}}{\partial x} + \frac{\partial H\bar{v}^2}{\partial y} + fH\bar{u} &= -gH\frac{\partial \eta}{\partial y} - \frac{gn^2\bar{v}\sqrt{\bar{u}^2 + \bar{v}^2}}{H^{1/3}}, \end{aligned} \tag{2.18}$$

where n is the Manning's coefficient. As we can note in Eqs. 2.18, the friction terms are important at shallow areas, when H becomes small. Therefore, it is

essential to properly specify the Manning coefficient according to the geometry of shallow and inundated areas. However, we have noticed that there is a debate about the adequate values for the Manning's coefficient for tsunami simulations, especially in dense urban areas, as described by Bricker et al. (2015). In this study we use mean values proposed for coastal and riverine areas (Bricker et al., 2015). The set of equations in Eq. 2.13 and Eq. 2.18 are known as the non linear shallow water wave equations. We remark here that wave frequency dispersion and breaking are not captured in the shallow water wave equations models. The equations can also be derived by adopting a spherical coordinate system (Wang, 2009), which are necessary for large domains (> 100 km) over the Earth's surface.

Two boundary conditions were used to derive the shallow water wave equations, at the bottom and the surface. To solve the surface elevation and velocity field over a domain we need to impose further boundary conditions at the boundaries of the domain. We describe here the two types of boundary conditions used in this thesis.

1. Radiation boundary conditions: This is known as open boundary or absorbing boundary, as well. It aims to completely transmit the wave outside the boundaries, suppressing any reflection. This boundary is used to limit the size of the computational domain, provided that tsunami waves which reflect with physical obstacles outside the domain do not influence tsunami solutions in assessed areas.

   Many models have been provided in the literature to simulate radiation boundary conditions with dfferent accuracies. Engquist and Majda (1977), for instance, presented a set of models with different levels of accuracy.

The simplest model adopts the method of characteristics and assumes linear shallow water waves. The radiation boundary condition for a one-dimensional propagation is given by,

$$\frac{\partial u}{\partial t} \pm \frac{\partial u}{\partial x} = 0, \tag{2.19}$$

where u is the wave celerity and x is the horizontal coordinate. This boundary condition precludes the reflection of waves coming from $+x$ (using $+\frac{\partial u}{\partial x}$) or $-x$ (using $-\frac{\partial u}{\partial x}$). A numerical finite difference implementation for the two dimensional version of Eq. 2.19 was derived by Liu et al. (1995), which can be used in shallow water wave models.

2. Moving boundary: This type of boundary is implemented when the boundary location changes due to the variation of the surface elevation in time. This is the case of the inundation process in coastal areas. As the water elevation increases, water will inundate previously dry areas. In the same way, a decreasing of the water elevation will dry previously wet areas. At any time the following set of equations are given at the shoreline,

$$\begin{aligned} \eta + b = 0 \quad \forall x, y \in \partial D_s(t), \\ \bar{u}_{(x,y,t)} = \bar{v}_{(x,y,t)} = 0 \quad \forall x, y \in \partial D_s(t), \end{aligned} \tag{2.20}$$

where $\partial D_s(t)$ is the shoreline and $-b$ is the vertical position of the seafloor with respect to the still water level (see Fig.2.4). This set of equations is computed at every time of the tsunami simulation, and the shoreline position is updated. A numerical finite difference implementation is presented in Liu et al. (1995). This implementation is adopted in the COMCOT numerical model.

Analytical solutions for simplified geometries

In this thesis we use analytical solutions of the shallow water wave equations to illustrate a simplified experiment in chapter 3, which assess the sensitivity of the tsunami to earthquake characteristics. Analytical solutions for the tsunami shoreline elevation and velocity, in simplified coastal geometries, have been extensively studied in the literature (Synolakis, 1987; Kanoglu, 2004; Tinti and Tonini, 2005; Madsen and Schaeffer, 2010; Sepúlveda and Liu, 2016). These solutions are mainly based on the hodograph transformation proposed by Carrier and Greenspan (1958).

Sepúlveda and Liu (2016) proposed analytical solutions for two beach geometries and arbitrary shapes of the tsunami initial condition. In this thesis we consider the beach geometry of Fig. 2.5, in which a tsunami is generated in a constant depth region. Then, tsunami waves propagate through the beach with slope γ, perturbing the shoreline elevation Se and shoreline velocity U. The runup R_u is defined as the maximum Se. The analytical solution assumes that tsunami waves are initially outside the sloping beach and they are linear in the constant depth region. Thus, the tsunami shoreline elevation Se and velocity U are expressed by,

$$Se(\tau) = \sum_{n=1}^{N} \frac{\sqrt{2h}\pi}{(L_n\gamma)^{1/2}} A_n \cos\left[\frac{2\pi\sqrt{gh}}{L_n}\left(\tau - t_n - 2\frac{h/\gamma}{\sqrt{gh}}\right) + \frac{\pi}{4}\right] - \frac{U(\tau)^2}{2g}, \tag{2.21}$$

$$U(\tau) = \sum_{n=1}^{N} \frac{2\pi^2\sqrt{2g}h}{(L_n\gamma)^{3/2}} A_n \sin\left[\frac{2\pi\sqrt{gh}}{L_n}\left(\tau - t_n - 2\frac{h/\gamma}{\sqrt{gh}}\right) + \frac{\pi}{4}\right], \tag{2.22}$$

where g is the gravity acceleration, and A_n, L_n and t_n are the amplitude, wavelength and phase shift of the $n-th$ tsunami initial condition Fourier component, respectively. h is the water depth in the constant depth region and τ is an non-

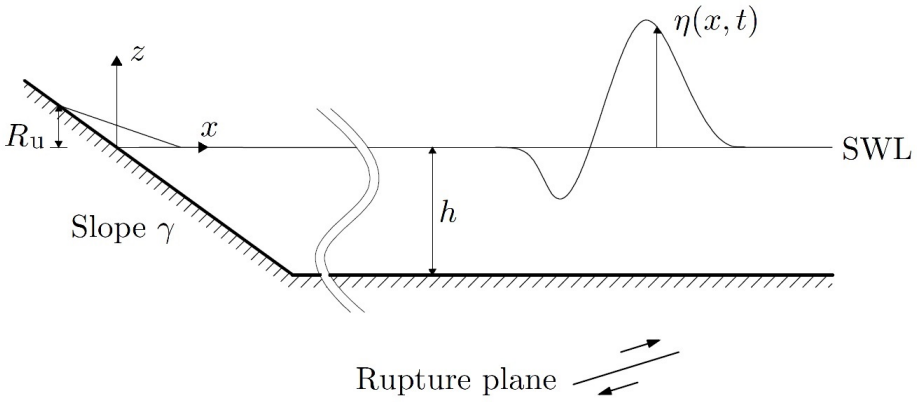


Figure 2.5: Incident wave propagating in the negative x direction over a constant depth ocean followed by a uniform beach slope. SWL denotes the Still Water Level.

physical coordinate related with the time and shoreline velocity as,

$$t = \frac{U}{g\gamma} + \tau. \tag{2.23}$$

We remark that the solutions presented here do not consider wave breaking and frequency dispersion.

Numerical model COMCOT

In the application cases of this study we use the COMCOT (Cornell Multi-grid Coupled Tsunami model) tsunami model. COMCOT adopts a explicit staggered leap-frog finite difference schemes to solve the nonlinear shallow water wave equations (Wang, 2009). The model solves the equations with an accuracy of the order $O(\Delta x^2, \Delta t^2)$, where Δx is the grid size and Δt is the time step. Due to stability considerations of the numerical scheme, Δx and Δt must fulfill the

Courant criterion,

$$\Delta t < \frac{\Delta x}{\sqrt{gH}}, \tag{2.24}$$

where $\sqrt{gH}$ is the wave celerity. It is important to mention that COMCOT automatically specifies $\Delta t \leq 0.5 \frac{\Delta x}{\sqrt{gH}}$ to guarantee numerical stability.

We remark here that wave breaking is not completely captured in the shallow water wave equations models, including COMCOT. However, the Manning quadratic law is used to model bottom friction, which can also be adopted to model the breaking wave induced energy dissipation. It is also important to remark that radiation and moving boundary conditions are implemented numerically in COMCOT to simulate the open boundaries and the shoreline movements, respectively. Further details of the numerical implementation can be revised in Wang (2009). COMCOT outputs are the surface elevation η and the fluxes $H\bar{u}$ and $H\bar{v}$, from where $\bar{u}$ and $\bar{v}$ can be derived.

2.3 Tsunami hazard assessment

In this section we describe the relevant aspects of the tsunami hazard, which are commonly associated with human loss and property damage. Then, we describe the tsunami hazard assessment approaches.

The Intergovernmental Oceanographic Commission of UNESCO presented a guide to survey tsunami effects at the coast (IOC, 1998), with emphasis on life loss and infraestructure damage. This guide identifies primary and secondary agents of these effects. The primary agents are the hydrostatic (pressure and bouyancy) and hydrodynamic forces (surge and drag due to currents). Fig. 2.6 presents relevant parameters related with these forces. The runup is the vertical

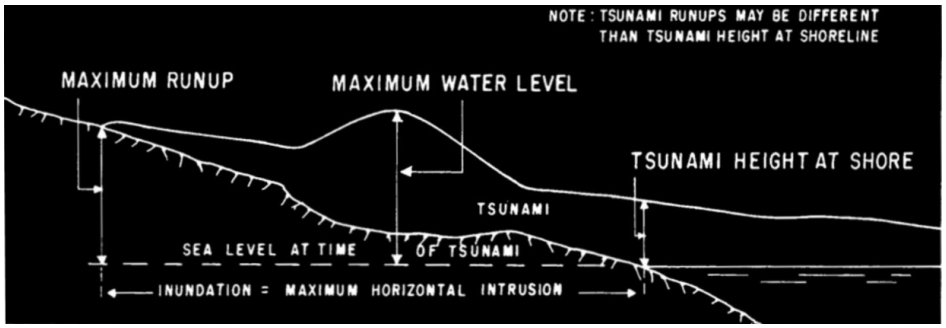


Figure 2.6: Relevant heights of tsunamis (IOC, 1998).

height that tsunami waves reach on land with respect to the still water level (at the time of the tsunami). The area between the still water shoreline (at the time of the tsunami) and the maximum runup line is defined as the inundation area. Fig. 2.7 shows a plan view of the inundation area at Juan Fernandez Island, Chile, due to the 2010 Chilean earthquake and tsunami (Winckler et al., 2010). Further metrics can be also assessed, such as the maximum tsunami amplitude, defined as the vertical distance between the still water level to the maximum tsunami elevation at a given location, or tsunami maximum velocities, which are associated with drag forces on structures and the transport of debris. The secondary agents, on the other hand, are the consequence of primary agents, such as the impact by debris, fires, explosions, contamination from hazardous materials, toxic fume releases, scouring of structure foundations, among others. The assessment of secondary agents is complex and usually requires additional models and information. In this study we only consider the primary agents. Specifically, we consider the maximum tsunami amplitudes, the runup and the inundation.

As we mentioned in the introduction, there are two typical approaches for assessing tsunami hazard, the so-called worst case scenario assessment and the

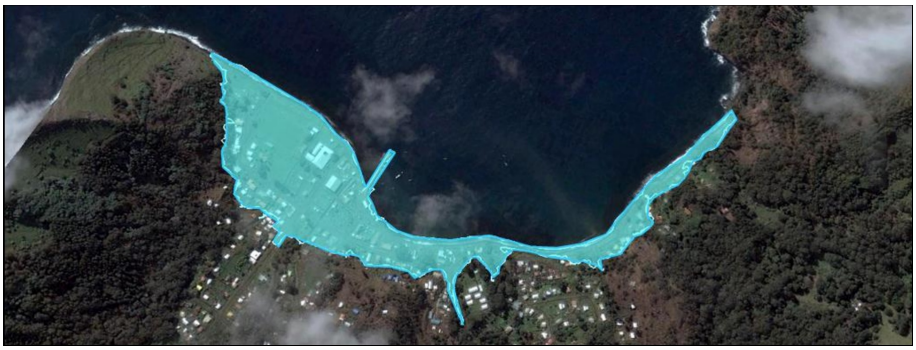


Figure 2.7: Surveyed inundation of the 2010 tsunami in Juan Fernandez Island, Chile (Winckler et al., 2010).

probabilistic tsunami hazard assessments (PTHA). While the worst case scenario assessment analyzes a single tsunami scenario from a specific seismogenic region and due to a specific earthquake magnitude, the PTHA is rather complex and considers many seismogenic regions and earthquake magnitudes. Geist and Parsons (2006) presented the framework for PTHA, which may include uncertain inputs as follows. Consider earthquakes of magnitude within the interval $M_{Wj} \pm \frac{\Delta M_{Wj}}{2}$, occurring in the seismogenic region x_i. The probability for having such an earthquake magnitudes in x_i after a time T from today, can be modeled as a Poisson process,

$$P\left(t > T | M_{Wj} \pm \frac{\Delta M_{Wj}}{2}, x_i\right) = e^{\left(-\lambda^{EQ}_{(M_{Wj}, x_i)} T\right)}, \tag{2.25}$$

where $\lambda^{EQ}_{(M_{Wj}, x_i)}$ is the mean arrival rate of earthquakes of magnitude within $M_{Wj} \pm \frac{\Delta M_{Wj}}{2}$ in x_i, which can be estimated by means of the Gutenberg-Richter (G-R) law (e.g. Shearer (2009)). The G-R law states that the recurrences of different earthquake magnitude intervals can be related. Thus, the recurrence of large earthquakes can be inferred from the statistics of small earthquakes within x_i. An example of the G-R law is provided in Chapter 5. Consider now that we are interested on the exceedance of a certain tsunami metric value h_{crit} at a given

assessed location. Due to the uncertainty of input data, earthquakes with magnitudes $M_{Wj} \pm \frac{\Delta M_{Wj}}{2}$ in x_i have a chance $P_h(h > h_{crit}|M_{Wj} \pm \frac{\Delta M_{Wj}}{2}, x_i)$ to exceed the tsunami metric value h_{crit}. By properties of Poisson processes (e.g. splitting Poisson processes in Ross (2014)), the probability to exceed h_{crit} after a time T from today, given earthquake magnitudes within $M_{Wj} \pm \frac{\Delta M_{Wj}}{2}$ in x_i is given by,

$$P_h\left(h > h_{crit}, t > T|M_{Wj} \pm \frac{\Delta M_{Wj}}{2}, x_i\right) = e^{\left(-\lambda^{EQ}_{M_{Wj},x_i} P_h(h>h_{crit}|M_{Wj}\pm\frac{\Delta M_{Wj}}{2},x_i)T\right)}. \tag{2.26}$$

If we further assume that earthquakes in different magnitude intervals $M_{Wj} \pm \frac{\Delta M_{Wj}}{2}$ and in different seismogenic regions are independent, we can obtain the joint probability to exceed h_{crit} after a time T (Ross, 2014),

$$P_h\left(h > h_{crit}, t > T\right) = \prod_j \prod_i e^{\left(-\lambda^{EQ}_{M_{Wj},x_i} P_h(h>h_{crit}|M_{Wj}\pm\frac{\Delta M_{Wj}}{2},x_i)T\right)} = e^{-\frac{T}{T_R(h_{crit})}}, \tag{2.27}$$

where the products consider all the earthquake magnitude intervals and seismogenic regions generating tsunamis in the assessed location and T_R is the joint return period exceeding h_{crit}. Note in Eq. 2.27 that T_R is equal to,

$$T_R(h_{crit}) = \frac{1}{\sum_j \sum_i \lambda^{EQ}_{M_{Wj},x_i} P_h(h > h_{crit}|M_{Wj} \pm \frac{\Delta M_{Wj}}{2}, x_i)}. \tag{2.28}$$

Finally, note that the probability to exceed h_{crit} at least one time before T, is equal to,

$$P_h(h > h_{crit}, t \leq T) = 1 - P_h(h > h_{crit}, t > T). \tag{2.29}$$

We remark that the determination of $P_h\left(h > h_{crit}, t > T|M_{Wj} \pm \frac{\Delta M_{Wj}}{2}, x_i\right)$ in Eq. 2.26 to Eq. 2.28 is not straightforward. As a consequence, the uncertainties have rarely been included or have been addressed inconsistently in existing methodologies. In section 2.4 we describe these methodologies in detail. Then, we propose a new consistent methodology to incorporate uncertain inputs for both, the worst case scenario assessment and the PTHA approaches.

2.4 Random elements and stochastic model

The incorporation of uncertain inputs in tsunami hazard assessments is complex because they do not have a unique value. Thus, these inputs and their corresponding tsunami model results are not deterministic. One alternative is to solve this problem by adopting a stochastic approach, in which we model inputs and outputs as random elements. The tsunami model, on the other hand, takes the form of a stochastic model.

2.4.1 Random elements

A random element is a function whose output is a $\mathbb{R}$-valued variable (random value), a $\mathbb{R}$-value vector (random vector), a real-valued continuous function defined in $I = [0, t]$ (stochastic process) or a real-valued continuous function defined on a subset $D \in \mathbb{R}^d, d > 1$ (random field). The output of the random element is so-called sample, and is taken randomly from a set of possible outputs, ruled by a probability distribution.

Formally, the random element is a function from a probability space to a measurable space. The probability space is defined by a sample space Ω, a σ-field $\mathscr{F}$ and a probability measure P. The sample space Ω is defined as the set of all possible elementary outcomes that may occur in a particular experiment. The outcomes can be numeric or not. The σ-field $\mathscr{F}$ is conformed by a collection of outcome subsets of interest, contained in Ω, which gives meaningful statements for a particular analysis. The σ-field $\mathscr{F}$ must (1) contain the empty set, (2) if $A \in \mathscr{F}$, then $A^c \in \mathscr{F}$ and (3) the union of partitions belonging to the

$\mathscr{F}$ must belong to $\mathscr{F}$. The probability measure is defined as a function with (1) $P(\varnothing)$=0 and (2) $\sum_{i=1}^{\infty} P(A_i) = 1$ for mutually disjoint sets belonging to the σ-field. That is, the probability measure quantifies the chance to obtain a possible event (or set of events). A measurable space, on the other hand, is conformed by a sample space S_S and a Borel σ-field $\mathscr{S}$. The sample space S_S is a metric space, conformed by a set of outcomes where function D exists, having the following properties: (1) it is positive definite, (2) obey symmetry and (3) follows the triangle inequality. This means that function D is a measure of distance defined in all S_S. The Borel σ-field, $\mathscr{S}$, is conformed by sets of intervals belonging to S_S. A random element X is, therefore, a function $X : \mathscr{F} \to \mathscr{S}$, provided that the events in $\mathscr{F}$ are mapped into $\mathscr{S}$ in a one-to-one relation and $X^{-1}(\mathscr{S}) \subset \mathscr{F}$. Random elements have mathematical meaning, allowing the computation of some measures such as the expectation, variance and other moments.

As an example of these definitions, let us consider the random variable X defined as the tsunami height at a given assessed location due to earthquakes in a given seismogenic region. The sample spaces Ω and S_S may consider all the possible earthquakes in a given seismogenic region and the generated tsunami heights at the assessed location, respectively. The events of the Borel σ-field $\mathscr{S}$ of S_S can be defined, for instance, as intervals of tsunami height $\{[0, h_1), [h_1, h_2)...[h_n, \infty)\}$, with $h_1 > h_2 > ... > h_n > \infty$. The events of the σ-field $\mathscr{F}$ of Ω are then sets of earthquakes which corresponds to each event of $\mathscr{S}$. Finally, the probability measure P gives probability of occurrence of the defined events.

We remark that the definition of random elements as functions obeys relevant mathematical arguments. For instance, this definition allows the trans-

formation from one random element to another. In the particular case of this thesis, the earthquake slip and bathymetry are modeled as functions of a set of Gaussian random numbers (K-L coefficients).

Type of random elements

In this study we model uncertain inputs and outputs as random elements. The earthquake location and the maximum tsunami amplitude are modeled as random values. A random value is a function which gives a single number as an output. The earthquake slip distribution and bathymetry, on the other hand, are modeled as random fields. A random field is a function defined at the spatial coordinates $D \in \mathbb{R}^d, d > 1$. The values of the random field at different spatial coordinates may or may not be dependent (i.e. the value at one spatial location influences the value at another location). We remark here that the modeling of random fields is complex because they have an infinite random dimension (i.e. they are defined by infinite random values corresponding to infinite spatial locations). In this thesis we construct random field samples by adopting an approximate approach. This approach requires samples of random vectors (i.e. vector with coefficients for the Karhunen-Loeve expansion in section 2.4.2). A random vector is defined by more than one random value, which are known as coordinates. These coordinates may or may not be dependent.

Probability distribution and moments

The representation of the values (or intervals) that a random element can take, in terms of their chance of occurrence, is called the probability distribution.

For the case of uncountable random elements (i.e. continuous random elements), the probability that a specific value occurs is zero. The probability of occurrence of an interval, however, is greater than zero. Consider a borel σ-field $\mathscr{S}$ formed by all possible $(a, b]$, $(-\infty, a]$, (b, ∞), $(-\infty, \infty)$, $\emptyset$, where a,b $\in \mathbb{R}^d$, $a < b$ and $d \geq 1$ is the dimension of the random element. The function F: $S_S \rightarrow [0, 1]$ defined by $F(x) = P(x \in (-\infty, x])$, where S_S is the sample space and x is a value of the random element, is called the cumulative density function (cdf). Notice that x may have a dimension greater than 1 (e.g. random vectors), or even infinite (e.g. random fields). The probability density function (pdf), on the other hand, is a non-negative function $f(x)$ defined as,

$$f(x) = \frac{\partial^d F}{\partial x_1 \partial x_2 ... \partial x_d}. \tag{2.30}$$

An additional function is the probability marginal distribution F_j for multi-dimensional random elements, which informs about the probability of occurrence of an interval in the dimension j, regardless of the values in other dimensions,

$$F_j(x_j) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \int_{-\infty}^{x_j} \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f\left(x_1, ..., x_{j-1}, y, x_{j+1}, ..., x_d\right) dx_1 ... dx_{j-1} dy dx_{j+1} ... dx_d. \tag{2.31}$$

Note that we can also obtain the marginal pdf $f_j(x_j)$ of x_j. From the probability distributions, presented above, we can compute the moments, which inform about properties of the random elements and their distributions. The moments are defined as,

$$\mu(q_1, q_2, ..., q_d) = E\left[\prod_{j=1}^{d} (x_j - \mu_j)^{q_j}\right], \tag{2.32}$$

where μ_j corresponds to the mean of the $j-th$ random dimension and it is defined as,

$$\mu_j = \int_{-\infty}^{\infty} x_j f_j(x_j) dx_j, \tag{2.33}$$

The expectation operator $E[]$ is defined as,

$$E[A(x)] = \int_I A(x) f(x) dx, \tag{2.34}$$

where I is the sample space and $A(x)$ is any function of the random element x.

The determination of the pdf's and cdf's of random vectors and random fields is rather complex because they are multidimensional. The cdf and pdf of some random vectors and fields, although, can be completely defined by their probability marginal distribution and covariance. The covariance between the $i-th$ and $j-th$ locations, $c(x_i, x_j)$, is given by:

$$c(x_i, x_j) = \mu(q_i = 1, q_j = 1) = E\left[(x_i - \mu_i)(x_j - \mu_j)\right]. \tag{2.35}$$

For simplicity, the existing methodologies reviewed in section 2.5 and this thesis consider random fields which are completely defined by the probability marginal distribution and covariance.

Random field models can be further simplified be adopting homogeneous probability properties. Homogeneous random fields have a probability marginal distribution and a covariance function which do not depend on the position in the field. As a consequence, the covariance function is simply expressed in terms of the spatial distance between any $i-th$ and $j-th$ locations, $\xi = x_i - x_j$. In chapter 3 we adopt a homogeneous random field to model the uncertain earthquake slip distribution. Further justifications and details are provided in the state of the art review in section 2.5.1.

It is also important to mention that some random fields are not entirely uncertain and some positions have known values. These fields are called conditional random fields. In these fields the covariance and the marginal distribution vary spatially and depend on the known values.

To construct a conditional random field, we first set the priori random field which is not conditional to known data. This priori random field is homogeneous with known covariance c_p and marginal distribution with mean μ_p. The conditional random field, conditional to a set of known values z at some field positions, is defined by a conditional covariance c_b and conditional mean μ_b (Kelker, 1970),

$$c_b(x_i, x_j) = c_{p11} - c_{p12}(c_{p22})^{-1}c_{p21}, \tag{2.36}$$

$$\mu_b = \mu_{p1} + c_{p12}(c_{p22})^{-1}(z - \mu_{p2}), \tag{2.37}$$

where the matrices c_{p11}, c_{p22}, c_{p12} and c_{p21} are blocks of c_p. c_{p11} are the covariances of the positions with uncertain values, c_{p22} are the covariances of the positions with known values and c_{p12} are the covariances between positions with uncertain values (rows) and positions with known values (columns). Finally, $c_{p12} = c_{p21}^T$. The vector μ_{p1} contains the mean of field positions which have uncertain values, while μ_{p2} contains the mean of the positions with known values. Note that the size of $c_b(x_i, x_j)$ and μ_b is smaller than the size of c_p and μ_p, since the former only consider positions with uncertain values. We remark that $c_b(x_i, x_j)$ and μ_b vary spatially and, thus, conditional random fields are not homogeneous. The bathymetry field model of Chapter 4 is modeled as a conditional random field, which is conditional to known bathymetry data at some geographical positions. Further justifications and details are provided in the state of the art review in section 2.5.2.

2.4.2 Sampling of random elements

The generation of samples of random vectors with low dimension is straightforward and can be computed by the many toolboxes available in programming languages. The generation of samples of random fields, conversely, is rather complex because these random elements have infinite random dimensions. However, parametric model can be adopted, which model random fields as a sum of deterministic function modes multiplied by random values. The random fields are then represented by random vectors with finite dimensions.

Random fields with finite variance can be modeled by a Karhunen-Loeve (K-L) expansion (Grigoriu, 2012). Unlike the Gaussian random field, the K-L expansion of non-Gaussian random fields is very complex and sometimes unfeasible. To overcome this difficulty, a K-L expansion is first adopted to generate samples of a Gaussian field $G(x)$, in which $x \in \mathbb{R}^2$ denotes position in the random field. Then, samples of $G(x)$ are used in a translation model (Grigoriu, 1995) to build samples of any non-Gaussian random field $S(x)$. Following are technical details on our strategy for generating random field samples.

K-L expansion

Let $G(x)$, $x \in \mathbb{R}^2$, be a real-valued homogeneous Gaussian field with mean $\mu = E[G(x)]$, variance $\sigma^2 = E[(G(x) - \mu)^2]$, and covariance function $c_G(\xi) = E[(G(x + \xi) - \mu)(G(x) - \mu)]$, $x, \xi \in \mathbb{R}^2$. The field admits the representation,

$$G(x) = \mu + \sum_{k=1}^{\infty} \lambda_k^{1/2} Z_k \psi_k(x), \tag{2.38}$$

where $\{Z_k\}$ are independent Gaussian variables with mean 0 and variance 1 and $\{\lambda_k, \psi_k(x)\}$ are the eigenvalues and eigenfunctions of the covariance function of

$G(x)$, i.e., they solve the eigenvalue problem,

$$\int_{\mathbb{R}^2} c_G(x-y)\,\psi(y)\,dy = \lambda\,\psi(x), \quad x \in \mathbb{R}^2. \tag{2.39}$$

We note that the equality in Eq. 2.38 holds in the mean square sense and the representation of $G(x)$ in this equation is referred to as K-L expansion. It is important to mention that the solution of Eq. 2.39 is not straightforward. However, approximate numerical methods can be adopted. In this study we use the Nystrom method (Betz et al., 2014), which discretize the random field in a grid and transforms Eq. 2.39 into an alegrabraic equation. This method is associated with a discretization error, ϵ_d, which is described later.

The generation of samples of $G(x)$ involves three steps. First, the infinite series in Eq. 2.38 is truncated. It is common to retain the dominant n terms of this series, i.e., the terms corresponding to the largest n eigenvalues. This truncation is associated with an error, ϵ_t, which is described later. Accordingly, the field $G(x)$ is approximated by,

$$G_n(x) = \mu + \sum_{k=1}^{n} \lambda_k^{1/2}\, Z_k\, \psi_k(x), \tag{2.40}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_n$ are the top n eigenvalues. Second, samples of the random variables $\{Z_k\}$ are generated. For example, the MATLAB function **randn** can be used to produce samples of this vector. Third, the samples of $\{Z_k\}$ are introduced in Eq. 2.40 to obtain samples of $G_n(x)$. It can be shown that $G_n(x)$ converges to $G(x)$ in mean square sense as $n \to \infty$, so that samples of $G_n(x)$ can be used as a substitute for samples of $G(x)$, provided n is sufficiently large. Note that the K-L expansion generates random field samples by means of a random vector, whose coordinates are the random coefficient of each K-L term.

Translation model

Non-Gaussian random fields $S(x)$ can be described by a translation model, i.e.

$$S(x) = F^{-1} \circ \Phi(G(x)), \quad x \in \mathbb{R}^2, \tag{2.41}$$

where F is an arbitrary cdf and Φ denotes the Gaussian cdf. The random field $S(x)$ is homogeneous and has two notable properties. First, its marginal distribution is F since, $P(S(x) \leq \alpha) = P(F^{-1} \circ \Phi(G(x)) \leq \alpha) = P(\Phi(G(x)) \leq F(\alpha)) = F(\alpha)$, by properties of F and Φ. Second, the covariance functions of $S(x)$ and $G(x)$ are rather similar. Generally, it is possible to find a covariance function of $G(x)$ such that the covariance function of $S(x)$ is close to a target covariance (Grigoriu, 1998). Thus, samples of $S(x)$ result directly from samples of $G(x)$ and the definition of the random field in Eq. 2.41. Examples of translation random functions can be found in Grigoriu (1995, 1998).

2.4.3 Stochastic models for uncertainty propagation

The uncertainty propagation is a stochastic model which relates the uncertain inputs and outputs. The Monte Carlo simulation is the most used model of this kind, where a large set of samples are randomly selected and used to run a deterministic model several times, which could be, for instance, a typical deterministic tsunami simulation model. Estimates of the mean, variance and higher moments of the output probability distribution are then obtained. The method is robust and simple but computationally expensive since it requires a large amount of samples for the convergence of model output statistics. This is especially true for multi-dimensional random elements and non linear models.

Therefore, other stochastic models were investigated in the framework of this thesis.

Alternatives to Monte Carlo simulations are the spectral models, such as the Polynomial Chaos Expansion (Constantine, 2007). The approach is similar to the Karhunen-Loeve Expansion presented above, where inputs and outputs are represented by deterministic basis functions (modes) multiplied by random values. This method is often much faster than Monte Carlo simulations for low-dimensional random elements inputs and linear models. However the method is intrusive, which means that the equations defining the relation between inputs and outputs is modified. This is a relevant issue since any change in the characteristics of the inputs may modify the equations of the stochastic model. A non-intrusive version has been proposed in the literature, but its accuracy is reduced. Another relevant issue is that the convergence of the statistics of random outputs reaches up to the second moments only (i.e. mean, variance, covariance). Therefore, the method is not capable to estimate small probabilities of exceedance, which are commonly defined by the tails of the probability distributions. Hence, the Polynomial Chaos Expansion is not suitable for this thesis.

Field et al. (2015) compared different types of stochastic models, including the spectral methods, collocation method and reduced order models. Based on this comparison, it is observed that Stochastic Reduced Order Model (SROM) (Grigoriu, 2009) performs relatively better in terms of computational cost and accuracy.

SROM of tsunami model inputs

In this thesis we use the SROM, proposed by Grigoriu (2009). Consider a random vector Z, whose components can be, for instance, random numbers or coefficients of a K-L expansion describing uncertain inputs. An SROM model $\tilde{Z}$ of Z is a random element with a finite number of members in the sample space, $\{\tilde{Z}_1, \tilde{Z}_2, \tilde{Z}_3 ... \tilde{Z}_m\}$, which are a subset of Z and whose joint probability properties are similar to those of Z in the context of an objective function. The cardinality of $\tilde{Z}$ is known as the SROM size m. Each member has a corresponding probability, $\vec{p} = \{p_1, p_2, p_3 ... p_m\}$, which are obtained by solving the following optimization problem,

$$\begin{aligned} &\min_p \left\{ e_t\left(\tilde{Z}, \vec{p}\right) \right\}, \\ &\text{with the constraint} \quad \sum_{k=1}^{m} p_k = 1. \end{aligned} \tag{2.42}$$

The objective function is expressed as,

$$e_t\left(\tilde{Z}, \vec{p}\right) = \sum_{u \geq 1} \alpha_u e_u\left(\tilde{Z}, \vec{p}\right), \tag{2.43}$$

where $e_u\left(\tilde{Z}, \vec{p}\right)$ measures discrepancies between distribution, moments, covariance or other probability properties associated with Z and $\tilde{Z}$. The parameters α_u are weighting factors used to control the importance of each type of discrepancy in the optimization problem and which are specified for each particular uncertainty propagation. Note that the distribution, moments, covariance and other probability properties depend on $\tilde{Z}$ and $\vec{p}$, and so the objective function. The objective function is designed to minimize discrepancies of relevant probability properties, by testing different members for $\tilde{Z}$ and different values of $\vec{p}$. The optimization problem can also consider the probability properties of transformed variables which are result of a mapping from Z. The optimization problem of these cases also depends on $\tilde{Z}$ and $\vec{p}$. For instance, the uncertainty propagation

of Chapter 3, considers an optimization which minimizes the discrepancies of the probability properties of Okada's solutions, which are functions of the K-L coefficients and locations defining the earthquake samples (i.e. Z are earthquake characteristics).

The optimization problem to find the optimal samples $\tilde{Z}$ and their probabilities $\vec{p}$ is complex and challenging, especially when Z is multi-dimensional. Grigoriu (2012) proposed a sub-optimal, but straightforward procedure, which solves an optimization problem in terms of the probabilities only. The procedure involves three steps. First, generate a set with a sufficiently large n number of samples of Z. Second, generate n_{set} sets of m samples ($m << n$), randomly selected from the set of n samples. Third, consider each of the n_{set} sets to solve a optimization problem in terms of the probabilities only, by minimizing the objective function of Eq. 2.42. For example, the MATLAB function **fmincon** with an interior point algorithm can be used for this purpose. Finally, select the set of m samples, among the n_{set} sets, whose discrepancies are the smallest. Grigoriu (2012) demonstrated that the accuracy of the SROM, constructed with the sub-optimal procedure, increases with m and n_{set}.

2.5 State of the art quantifying the tsunami assessment uncertainties due to uncertain inputs.

From the literature review we noticed that the earthquake uncertainties have been addressed by using stochastic frameworks and by building earthquakes samples, known as synthetic earthquakes. The existing methodologies (e.g. Løvholt et al. (2012); Goda et al. (2014); Davies et al. (2015); Li et al. (2016);

De Risi and Goda (2016)), though, present inconsistencies in the generation of samples and in the propagation of uncertainties to tsunami assessments. The bathymetry uncertainties, on the other hand, has not been completely addressed so far, except for a few studies which have performed sensitivity analysis and made some attempts to generate synthetic bathymetry samples. In this section we review the cutting edge methodologies and their main issues.

The flow chart of Fig. 1.2 present the three stages that are adopted in stochastic approaches based on samples, which are used to assess tsunamis with uncertain inputs. Each stage faces significant challenges in the uncertain earthquake and uncertain bathymetry problems, which are discussed here.

2.5.1 Propagation of earthquake uncertainties

In this revision we focus on the aleatory uncertainty of future earthquakes. This uncertainty is not reduce with more data and is related with the assumed aleatory nature of the slip distribution and location of future events. We revise each stage of Fig. 1.2.

Probability properties of slip and location of earthquakes

The slip distribution has been conveniently modeled as an homogeneous random field in the literature. Homogeneous random field models require less information than non-homogeneous random fields because the probability properties do not vary spatially. The existing statistical information of the slip of past earthquakes, on the other hand, has been also obtained by assuming ho-

mogeneous statistical properties. It is important to remark that the relevance of this assumption has not been assessed so far. Further discussion is provided in section 6.1. As we mentioned above, the simplest random field models are characterized by two probability properties, the marginal distribution and the covariance. The determination of the true marginal distribution and covariance of slip distributions is challenging given the scarce information from past earthquakes. Gusev (2011) and Thingbaijam and Mai (2016) used the inversion solutions collected in the earthquake catalog SRCMOD (Mai and Thingbaijam, 2014) to construct the histogram of the slip in rupture areas of different seismogenic regions. The different histograms were then fitted into distribution functions (based on a modified Log-normal distribution in Gusev (2011) and based on an modified exponential distribution in Thingbaijam and Mai (2016)). By assuming that the slip random field is homogeneous and the histograms of the slip over the rupture area are similar to the slip marginal distribution, the proposed distribution functions can be adopted as marginal distribution for the slip random field. Some authors have also proposed simpler distribution functions, which have similar shape as the observed histograms. For instance, LeVeque et al. (2016) used the Log-normal distribution. It is important to remark that the construction of slip histograms from past earthquakes is challenging. These histograms are often associated with statistical errors, due to the small number of sub-faults used, and slip inversion errors. To illustrate this problem, Fig. 2.8 shows the estimated slip distribution of four recent earthquakes: the Japan 2011 (Hayes, 2011), Sumatra 2004 (Rhie et al., 2007), Chile 2010 (Hayes et al., 2013) and Chile 2014 (An et al., 2014). The slip has been obtained by using an inversion method based on seismic waves (Japan 2011, Chile 2010, Sumatra 2004) and tsunami records (Chile 2014). The blue bars on the right panels show the

histogram of the sub-fault slip values within an estimated rupture area, based on the scaling relation of Blaser et al. (2010) (black grids in left panels). Note that the small number of subfaults in all the events does not allow an accurate description of the true slip histograms. Moreover, it is not possible to determine if slip histograms of different seismogenic regions have a common shape or whether the Log-normal or exponential distribution better represents the slip marginal distribution.

The covariance of the slip distribution, on the other hand, was investigated by Andrews (1980), Herrero and Bernard (1994) and Mai and Beroza (2002). These authors computed the slip wave number spectra of past earthquakes in different seismogenic regions and observed that the slip power spectrum decays with a $k^{-\gamma}$ law for wave numbers greater than a characteristic value known as corner wave number. They also noticed that the parameter γ typically fluctuates around 2. The constant value of γ indicates that the slip has a fractal behavior (Andrews, 1980). Mai and Beroza (2002) and Raghukanth and Sangeetha (2016) used the obtained slip power spectra to test different covariance functions. To illustrate these tests, Fig. 2.9 shows the power spectrum of the slip of the 2008 Kashmir earthquake, obtained by means of a Fourier transform in Raghukanth and Sangeetha (2016). The Kashmir slip power spectrum is compared with the power spectra of the Gaussian, Exponential and Von Karman functions. As suggested by Mai and Beroza (2002) and Raghukanth and Sangeetha (2016), the Von Karman covariance function fitted the data best. The Von Karman covariance function has the form,

$$\begin{aligned} c(\xi) &= \sigma_s^2 \frac{2^{1-H_{VK}}}{\Gamma(H_{VK})} \bar{d}_\xi^{H_{VK}} K_{H_{VK}}(\bar{d}_\xi) \\ \bar{d}_\xi &= \sqrt{\frac{(\xi_x)^2}{a_x^2} - \frac{(\xi_y)^2}{a_y^2}} > 0, \end{aligned} \tag{2.44}$$

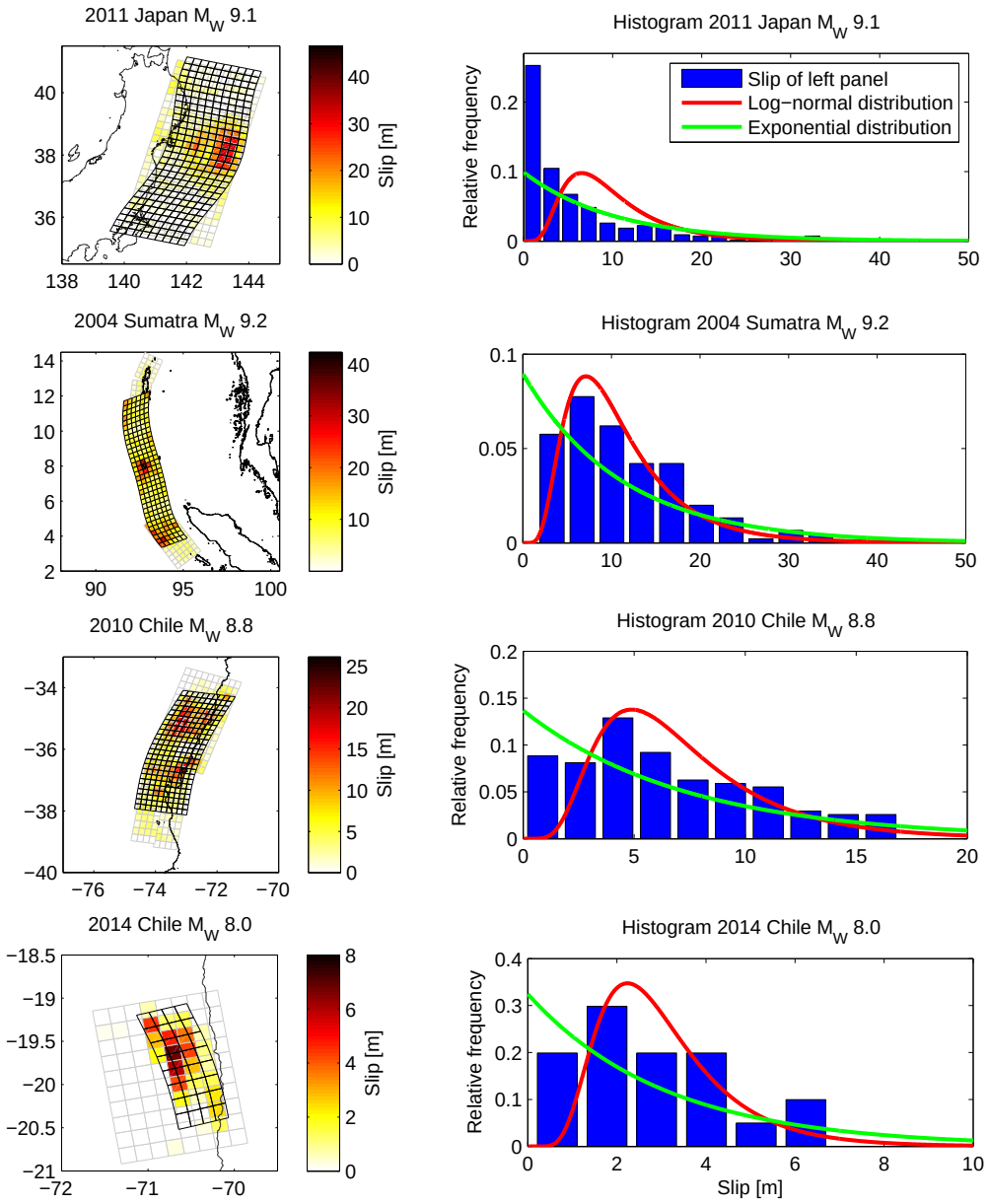


Figure 2.8: Left panels: Slip distribution of past earthquakes projected in the horizontal plane. The grey grid correspond to the discretization adopted in the inversion. The black grids correspond to the rupture area obtained with scaling relations (see Fig. 2.2). These latter grids follow the curvature of the trenches. Right panels: Slip histograms of the sub-faults inside the black grid of the left panel. Red curves correspond to the Log-normal marginal distribution, which is adopted in this study. Green curves correspond to the exponential marginal distribution. The rupture area of the 2004 Sumatra earthquake has been deformed to fit within the narrow seismogenic region.

where $\Gamma(H_{VK})$ is the Gamma function, $K_{H_{VK}}$ is the modified Bessel Function of the second kind of order H_{VK} and $\xi = (\xi_x, \xi_y)$ is the distance. The parameters a_y and a_x are the characteristic correlation lengths of the slip along the strike direction and dip direction, respectively. σ_s is the slip standard deviation. H_{VK} is known as Hurst number and is a shape parameter, which controls the decay at high wave numbers. Mai and Beroza (2002); Dorostian and Zaré (2009) and Raghukanth and Sangeetha (2016) have proposed scaling relations for the correlation lengths of the Von Karman covariance function, a_x and a_y, and the slip standard deviation, σ_s, in terms of the earthquake magnitude. In this study we adopted the scaling relations of Raghukanth and Sangeetha (2016) since they used large earthquakes ($M_W > 8$) to fit the scaling laws. The top panels of Fig. 2.10 shows the correlation lengths inferred by Raghukanth and Sangeetha (2016) and Mai and Beroza (2002) from past earthquakes events and the scaling relation of Raghukanth and Sangeetha (2016). The bottom panel of Fig. 2.10 also shows slip standard deviations of past events and the scaling curve of Raghukanth and Sangeetha (2016). We observe a significant scattering of data around the scaling relations of the correlation lengths. We think this scattering might be the consequence of some uncertainties in the calculation of slip power spectra, which often involves interpolations, Fourier transforms of slip in finite rupture areas and the use of slip inversion solutions (Raghukanth and Sangeetha, 2016; Mai and Beroza, 2002). Note that the inaccuracies of the slip power spectra, which might induce some errors in the estimation of correlation lengths, are evidenced in the example of Fig. 2.9. Further details of the calculation of slip power spectra is provided in section 3.3.2. Values for H_{VK} have also been proposed in the literature and usually ranging between 0 and 1 (Mai and Beroza, 2002). However, H_{VK} does not seem to be a relevant parameter for tsunami hazard purposes since

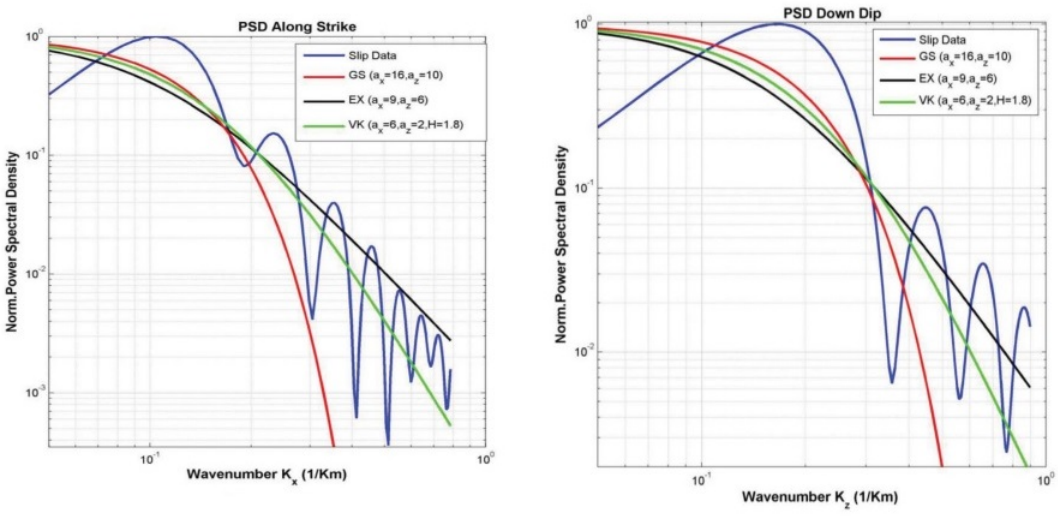


Figure 2.9: Example earthquake slip power spectrum, extracted from Raghukanth and Sangeetha (2016). The earthquake corresponds to the 2008 Kashmir event (M_W 7.6). Blue curves correspond to the one dimensional slip power spectrum, obtained with a Fourier transform. The red, black and green curves correspond to the Fourier transform of a Gaussian correlation function, an exponential correlation function and a Von Karman correlation function, respectively. The latter is adopted in this thesis.

tsunami responses are insensitive to the decay characteristics at high slip wave numbers, as discussed by Geist (2002) and Davies et al. (2015). From this review we conclude that the slip covariance function has been satisfactorily described in the literature. However, the various descriptions of the marginal distribution in different studies suggest that further and conclusive studies are required. In this study we assume a marginal distribution, which is described later.

The rupture area centroid location $\vec{x}_c$ can also be modeled as a random element. The probability laws of the location uncertainties, however, have not been addressed so far. In this study we assume a probability law for the location, which is described later.

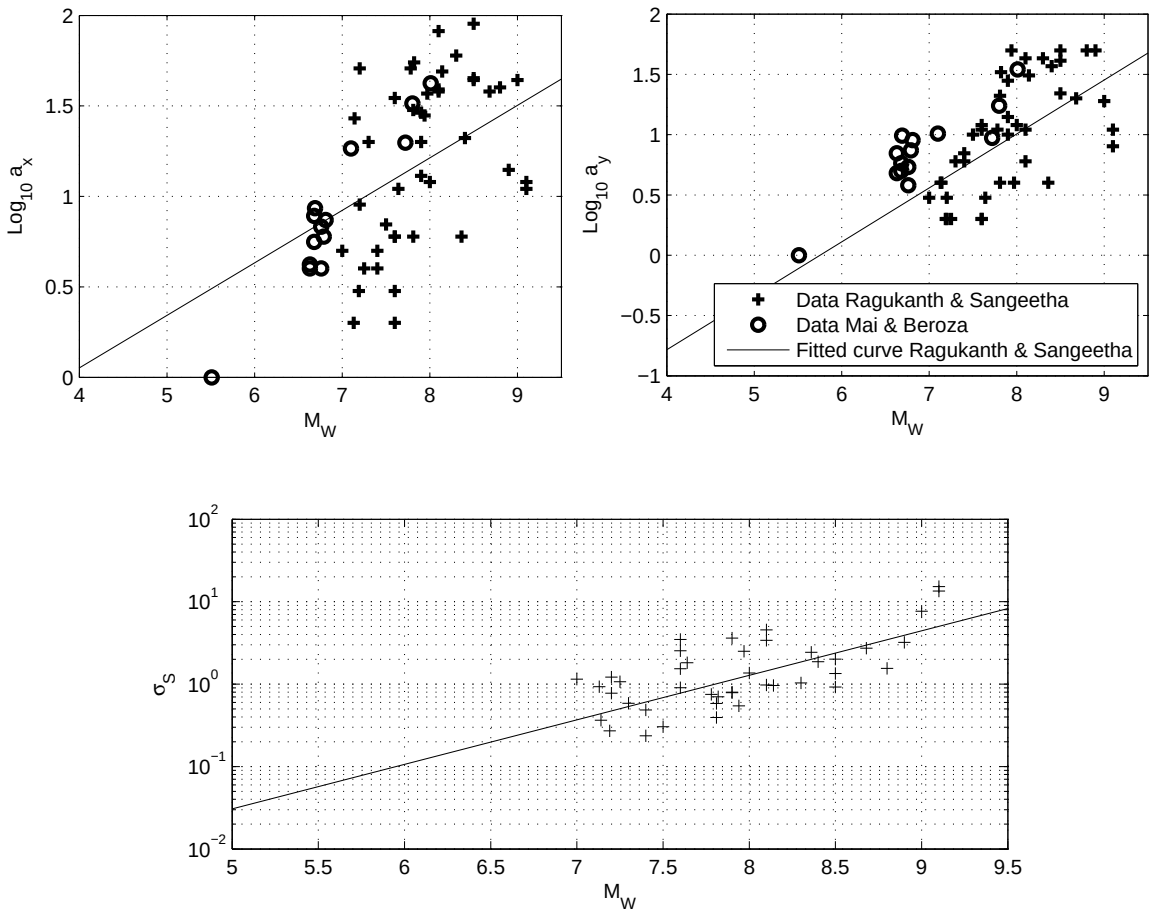


Figure 2.10: Top panels: Scaling relation curves for the correlation lengths a_x and a_y in meters. The curve were obtained from Raghukanth and Sangeetha (2016). The scatter data used by Raghukanth and Sangeetha (2016) and Mai and Beroza (2002) is also shown. Bottom panel: Slip standard deviation σ_s of past events (crosses) and scaling relation curve proposed by Raghukanth and Sangeetha (2016).

Generation of earthquakes samples

The existing methodologies for uncertainty assessments require the generation of samples of the slip distribution and location. Most of the existing studies (e.g. Løvholt et al. (2012); Goda et al. (2014); Davies et al. (2015); Li et al. (2016); De Risi and Goda (2016)) have obtained samples by using a Fourier transform with random phases (Herrero and Bernard, 1994; Gallovič and Brokešová, 2004b). Since the Fourier transform requires a rectangular rupture area, its application in large rupture areas with curved trenches requires further modifications. Li et al. (2016) generated samples by first adopting a rectangular rupture area and then deforming these samples to adapt them into a non-rectangular seismogenic region. The latter step, however, might lead to modifications of the deformed-samples covariance function. In this study, we utilize the K-L expansion. The K-L expansion allows the generation of samples for any non-rectangular geometry of the random field. We noted that LeVeque et al. (2016) were the first to adopt this approach for a random slip distribution sample generation. They used a pre-defined sub-fault discretization for the rupture area and a number of K-L terms, which are specified based on the accuracy of the seafloor deformation, obtained from the Okada model. In our proposed methodology we extend the application of the K-L expansion by considering an integral form of the eigenvalue decomposition, in which we can specify the K-L truncation and the sub-faults discretization in terms of the accuracy of the samples. The details will be provided later.

It is remarked here that most of the existing sample generation techniques adopt post-sampling procedures to control individual characteristics of the samples. For instance, the studies that considered Gaussian random fields replaced

negative values of the samples by zero or the absolute value so that the slip is non-negative everywhere (e.g. Gallovič and Brokešová (2004a); Herrero and Bernard (1994)). Others scaled the samples so the earthquake magnitude, associated with the slip, was fixed and equal to a desired value(e.g. Tselentis et al. (2006); Gallovič and Brokešová (2004a)). All these post-sampling procedure change the target marginal distribution and covariance function, originally specified for the random field. Another typical procedure is the selection of samples sharing some common features with a past earthquake, such as asperity locations (i.e., zone of maximum slip) (Goda et al., 2015). This practice modifies final probability properties as well. Davies et al. (2015) identified this problem and proposed an iterative method to preserve the random field covariance function and sample magnitude. Despite they preserved the target covariance, it is not clear if target marginal distributions were preserved. Further, this iterative approach could require a significant amount of time for a large number of samples. In our proposed methodology we avoid using any post-sampling procedure in the slip distribution samples. A relevant consequence of our proposed approach is that sample earthquake magnitudes are variable and deviate slightly from the expected magnitude specified for the random fields. In chapter 3 we shall explain the reason of this consequence and assess the severity of the magnitude variability.

Uncertainty propagation from earthquake uncertainties

The uncertainty propagation is also an important challenge for the uncertainty quantification in tsunami assessment (the step 3 in the stochastic approach shown in Fig. 1.2). Most of the studies have adopted Monte Carlo simulations

by which a set of earthquake samples are used to build a set of initial conditions for tsunami simulations. As LeVeque et al. (2016) also pointed out, the slip random field is multidimensional in the random space and, therefore, would require an unfeasible large amount of samples to properly perform a classic Monte Carlo. Recently, De Risi and Goda (2016) analyzed the convergence of tsunami statistics for an uncertainty assessment of a particular case. They used a Fourier Transform based approach to generate earthquake samples. From their results, we observe that Monte Carlo simulations with less than 200 samples are associated with inaccurate estimations of tsunami uncertainty statistics, especially those related with the tail of the distribution. Green's function techniques have been applied in some experiments in order to efficiently generate large number of samples (Li et al., 2016). These approaches, however, are applicable only to linear wave models and are not valid in regions where nonlinearity becomes important, such as the shallow or inundated areas. In this study we adopt the Stochastic Reduced Order Model (SROM), described above. The implementation of SROM for tsunami uncertainties will be further explained later.

2.5.2 Propagation of bathymetry uncertainties

As we mentioned in the introduction, tsunami models require bathymetry information in specific locations of a grid and the source of uncertainties are the errors of bathymetry data and interpolation errors. The bathymetry uncertainty has not been well examined in the literature, especially that related with interpolation errors. In this review we mention the most relevant studies.

The measurements of bathymetry and topography are based on different

technologies. The measurement errors can be classified in vertical and horizontal errors. The vertical errors is the difference between the true and measured depth at a given horizontal location. The horizontal error, on the other hand, is the difference between the measured and true horizontal location on the Earth. Table 2.1 shows the horizontal resolution of the bathymetry data sources used in this thesis. The accuracy of nautical charts varies with the location. According to Hare et al. (2011), the horizontal accuracy offshore (charts categorized as Zone of Confidence C) is about 500 m, while the vertical accuracy is 2m + 5% depth. The horizontal accuracy in ports and harbors (charts categorized as Zone of Confidence A1) is 5 m + 5% depth, while the vertical accuracy is 0.5 m + 1% depth. These latter type of nautical charts are very accurate. Their coverage, though, are limited to areas where navigation is relevant. Other bathymetry data, such as ship track soundings or private bathymetry surveys should be expected to have the same order of accuracy. The available nautical charts or any additional bathymetry information rarely cover the entire region where tsunamis are assessed. This is because they are rarely produced for international waters or at coastal regions where navigation is not relevant. In these regions, bathymetric models (known as digital elevation models, DEM) have been often adopted to estimate the bathymetry (e.g. An et al. (2014)). The most used DEM is the GEBCO dataset (Weatherall et al., 2015) which is based on nautical charts, ship-track soundings and a few survey campaigns. At unsurveyed regions GEBCO performs an interpolation complemented with satellite gravity measurements. Apparently, this interpolation lacks of a physical understanding of the bathymetry spatial characteristics. The accuracy of GEBCO data is expected to be the same as nautical charts at surveyed areas. The bathymetry accuracy at unsurveyed regions, though, is unknown and bathymetry data is

Database	Horizontal resolution
Nautical charts offshore	~20km
Nautical charts in ports	~20m
Gebco elevation model	30 arcsec (~1km)

Table 2.1: Bathymetry databases and resolution (Hare et al., 2011; Weatherall et al., 2015)

therefore unreliable.

Jakobsson et al. (2002), Calder and Mayer (2003) and Zambo et al. (2015) investigated the effect of errors of surveyed bathymetry data on interpolated grids for some site specific cases. For instance, Jakobsson et al. (2002) used classic Monte Carlo simulations to determine the uncertainty of interpolated grids within the Artic Ocean. They assumed that surveyed data have random and uncorrelated errors, following a Gaussian distribution. We observe two important issues with their adopted methodology. First, the probability properties of the random errors are not linked with any statistical analysis of errors in real surveyed data. Second, the methodology does not consider the uncertainty due to interpolation errors, but only the errors of surveyed data. Hence, these methodologies offer an incomplete description of bathymetry uncertainties. The uncertainty due to interpolation errors has only been investigated by sensitivity analyses, in which different interpolating techniques and high quality bathymetry data are adopted to assess the accuracy of interpolations in site-specific studies (Smith and Wessel, 1990; Erdogan, 2009). The conclusions of these latter studies, though, may not be applicable to other assessed areas. Since the accuracy of surveyed data (e.g. nautical charts in harbors) seems to be high and interpolation errors are still unknown, this thesis shall focus on the bathymetry uncertainty due to interpolation errors, which we expect to be dominant. To quantify the

bathymetry uncertainty we adopt a stochastic approach and follow the stages of Fig. 1.2.

The probability properties of the uncertain bathymetry due to interpolation errors (first stage of Fig. 1.2) have been scarcely addressed in the literature. However, we have found some relevant studies of the spatial statistics of surveyed bathymetry. During the sixties, Mandelbrot (1967) presented his seminal work defining fractals. His observations were focused on the shape of the surveyed coastline of England. After this study, many investigation were conducted to find the statistical properties of topo-bathymetry. Bell (1975), for instance, observed that the bathymetry power spectra in term of the wavenumber k are characterized by a constant decay which is proportional to $k^{-\gamma}$ for length scales smaller than one hundred kilometers. The parameter γ was estimated as 2. Based on this finding, Goff and Jordan (1988) were prompted to further describe all the statistics (marginal distributions, mean and covariance) of local bathymetry from echo-sounding based surveys. For simplicity, they adopted an homogeneous Gaussian random field. Williams et al. (2017), on the other hand, used a random field model to generate synthetic bathymetries of fjords, which were only random at small horizontal length scales. Large horizontal length scales, on the other hand, were deterministically specified. In this thesis we also propose to model uncertain bathymetry as a random field. From the mentioned studies, we have noted three relevant aspects regarding the modeling of the uncertain bathymetry. First, bathymetry uncertainty in unsurveyed areas are related with the natural spatial variation of bathymetry. Thus, statistical properties observed by Mandelbrot (1967) can be used to define the probability properties of the random field. Bathymetry uncertainty is not apparently bounded or biased. Therefore, we assume it is reasonable to adopt a Gaussian

marginal distribution (e.g. Goff and Jordan (1988)). The covariance, on the other hand, can be modeled by a Von Karman function, in which the spectral decay observed by Bell (1975) can be simulated. Second, large horizontal length scales of the bathymetry spatial variability (e.g. coastal slopes, trenches and mountain chains) can be known deterministically by using bathymetry data at surveyed locations (Williams et al., 2017). For instance, large horizontal length scales can be described by a linear interpolation using surveyed data. Third, bathymetry uncertainties are expected to depend on the distance to surveyed data. This latter aspect can be simulated by adopting a conditional random field model.

Since we have not identified studies addressing the bathymetry uncertainties, there is no investigation of the sample generation and uncertainty propagation methods (second and third stage in Fig. 1.2). In this study we propose a sample generation method using a K-L expansion. The uncertainty propagation is then performed by means of SROM. Further details are provided in Chapter 4.

CHAPTER 3

EARTHQUAKE MODEL

In this chapter we present a new method for the quantification of uncertainties in tsunami hazard assessments due to uncertain earthquakes. To demonstrate the relevance of earthquake uncertainties, we first present a motivational experiment. The new methodology is then presented and analyzed.

3.1 Motivation

We analyze the impact of a spatially varying slip distribution on the tsunami runup using analytical solutions for a simplified coastal geometry.

The solution of runup in terms of an earthquake with uniform slip distribution was investigated by Sepúlveda and Liu (2016). One of the problem setup, addressed in that reference, has the simplified coastal geometry of Fig. 2.5. The notation of the fault parameters is described in Fig. 2.1. We found that the non-dimensional runup $\frac{R_u}{s}\sqrt{\frac{\gamma d}{h}}$ has a non-monotonic relation with W/d, as it is shown in Fig. 3.1 for different dip angles δ. We identified three stages, a linear increase stage ($W/d < 1$ for all δ), a constant stage and a non linear increase stage. These stages are explained by the dependency of the runup on the amplitude and the characteristic wavelength of the initial condition. In the linear increase stage the seabed deformation increases in amplitude with a small increase of the characteristic wavelength. In the constant stage the amplitude and the wavelength increase by an amount that allows both effects to cancel each other. Finally in the non linear increase stage, the rupture plane approaches the seafloor and induces a high amplitude deformation and a short characteristic

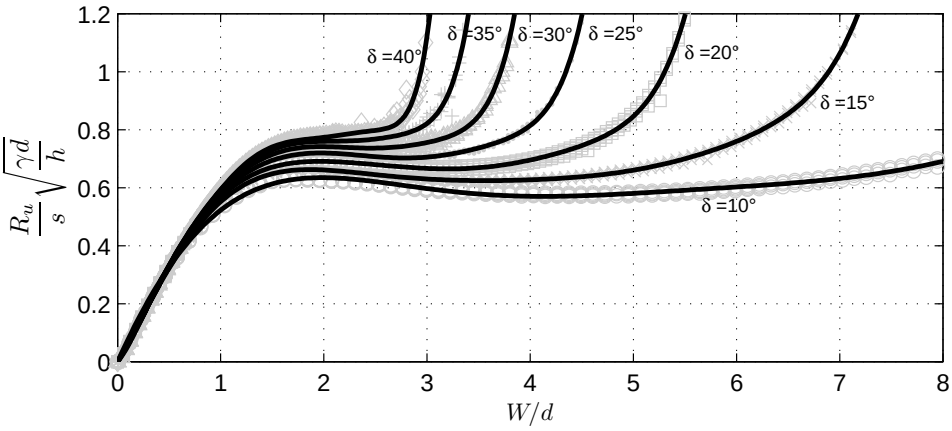


Figure 3.1: The dimensionless runup, $\frac{R_u}{s}\sqrt{\frac{\gamma d}{h}}$, for tsunami waves generated in a constant depth region, as a function of W/d and for different dip angles δ. Each symbol represents one numerical realization with a given set of parameters. Solid lines denote the 8^{th} degree polynomials fitting the data with same dip angle. $\diamond : \delta = 40°, + : \delta = 35°, \triangle : \delta = 30°, \bullet : \delta = 25°, \square : \delta = 20°, x : \delta = 15°, \circ : \delta = 10°$

wavelength. Thus, the runup increase until the wave breaks.

Here, we extend the analysis of Sepúlveda and Liu (2016) by considering a variable slip distribution in the rupture area. We specify different slip distributions where the average over W is preserved but the slip concentrates in segment of lengths $W/10$, $2W/5$, $7W/10$ and W at the upper and bottom edges of the fault. These segment lengths are similar to common slip correlation lengths for the Von Karman correlation function. Fig. 3.2 shows the results of runup for dip angles of 10^o, 20^o and 40^o, using the same non-dimensional parameters as Sepúlveda and Liu (2016). We observe that runups are considerably larger when the slip is concentrated at the top of the fault as compared with the case of a constant slip along the entire width. For most of the cases with the slip concentrated at the bottom edge the runup is also larger.

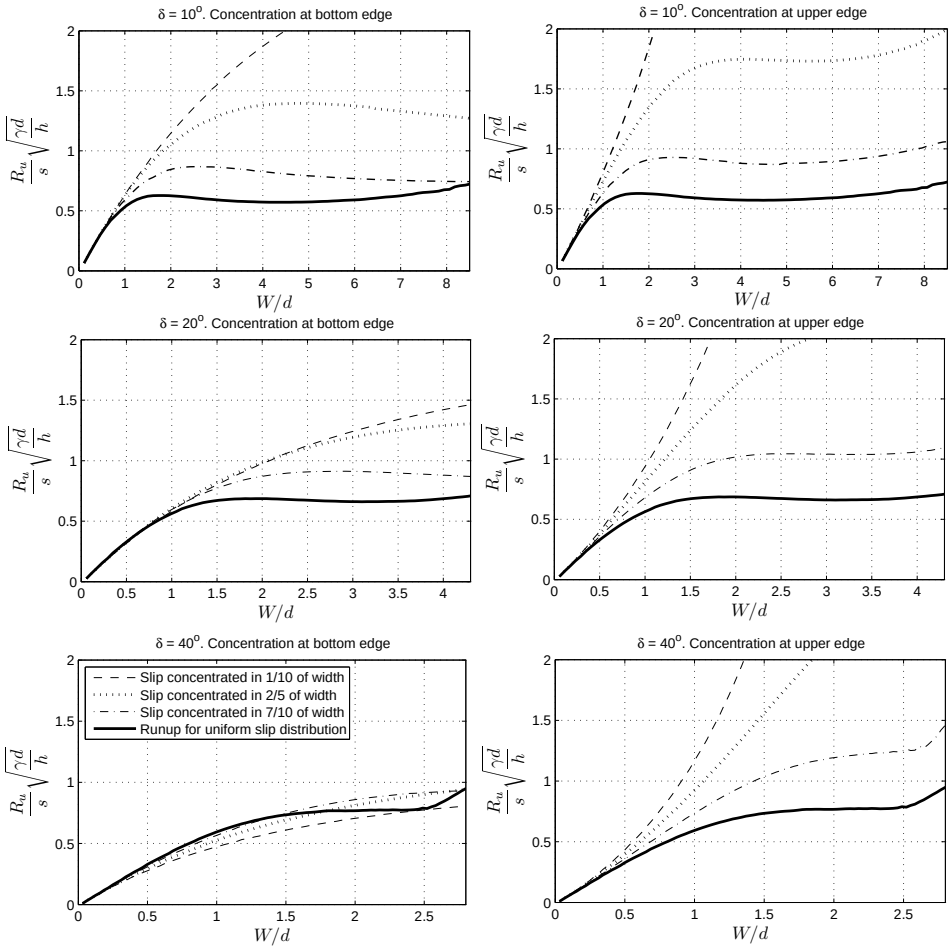


Figure 3.2: Runup values associated with slip which are concentrated in different rupture widths.

This motivational experiment demonstrates that the slip spatial variability and depth (or location) are relevant for the determination of the runup. The experiment also suggests that slip variability is often associated with larger runups. These observations have been also documented in previous studies (e.g. Løvholt et al. (2012)). Moreover, it has been also observed that it is more likely to estimate larger runups when variable slip distributions are adopted (Davies et al., 2015). It is important to mention that we only have analyzed the runup. However, velocities at the shore are also relevant parameters during the

inundation. Velocities are related with the temporal gradient of the shoreline elevation which are expected to increase with larger runups. Thus, it is expected to have similar observations. The analysis presented above concludes that the spatially variable slip and location are relevant parameters to assess the tsunami response accurately.

3.2 Earthquake uncertainty definition

We propose a new methodology for practical application cases. The methodology follows the flow chart of Fig. 1.2 and addresses the issues identified in the state of the art review section. To effectively explain each step of Fig. 1.2, we use an illustration case. We have chosen the site of the 2014 Chile earthquake (An et al., 2014; Lay et al., 2014; Gusman et al., 2015) since tsunami measurements were collected along the coast and at offshore DART stations (Meinig et al., 2005). Thus, we can compare these measurements with the results obtained by the present methodology. For the illustration case we have used a tsunami simulation model configured at the site (An et al., 2014). The 2014 Chilean earthquake has an estimated magnitude of M_W 8.0 (An et al., 2014; Gusman et al., 2015), which is used as the target earthquake magnitude to build hypothetical earthquakes. It is noteworthy to mention that slightly different magnitudes have been estimated for this event (e.g. M_W 8.1 in Yagi et al. (2014) and M_W 8.2 in Sipkin et al. (2000)). Hence, the target earthguake magnitude adopted in this illustration might differ from the true magnitude of the 2014 event. The geometry of the seismogenic region is non rectangular and it follows the curvature of the trench, as shown by the region bounded by red lines in Fig. 3.3. The minimum depth of the seismogenic region has been arbitrarily

specified as 1 km and the maximum depth has been specified as 55 km (Béjar-Pizarro et al., 2013; Métois et al., 2013). The dip angle within the seismogenic region is variable and has been obtained from the model of Tassara and Echaurren (2012). The averaged strike angle has been specified as 350^o and the rake angle has been conservatively specified as 90^o. The hypothetical earthquakes, studied here, have an uncertain distance to the trench but a fixed and known location along the trench. Since the dip angle is prescribed and the distance to the trench is uncertain, the fault depth is uncertain as well. This illustration case is referred to as the 2014 illustration case herein.

To model the slip distribution as a random field we define first the dimensions of the rupture area R (see Fig. 2.1), which are obtained from the scaling relations of Blaser et al. (2010), shown in Fig. 2.2. For the 2014 illustration case we consider a target earthquake magnitude of M_W=8.0, which corresponds to a fault width of W = 66.06 km and a fault length of L = 154.88 km.

The slip within R is then conveniently modeled as a homogeneous random field, according to the justification provided in the state of the art review (section 2.5.1). It is also assumed that the slip random field is completely defined by two properties, the marginal distribution and the covariance function. The marginal distribution can adopt one of the proposed distribution functions described in Section 2.5.1. For sake of simplicity, we adopt the Log-normal marginal distribution for the slip of the 2014 illustration case, which was also adopted by LeVeque et al. (2016). Two parameters of the Log-normal are required. They are the expected slip $\bar{S}$, and the slip standard deviation σ_s. To determine $\bar{S}$, the rupture area R (= $L \times W$) is inserted in the definition of the seismic moment in Eq. 2.1. The seimic moment of a M_W 8.0 earthquake is obtained with Eq. 2.2.

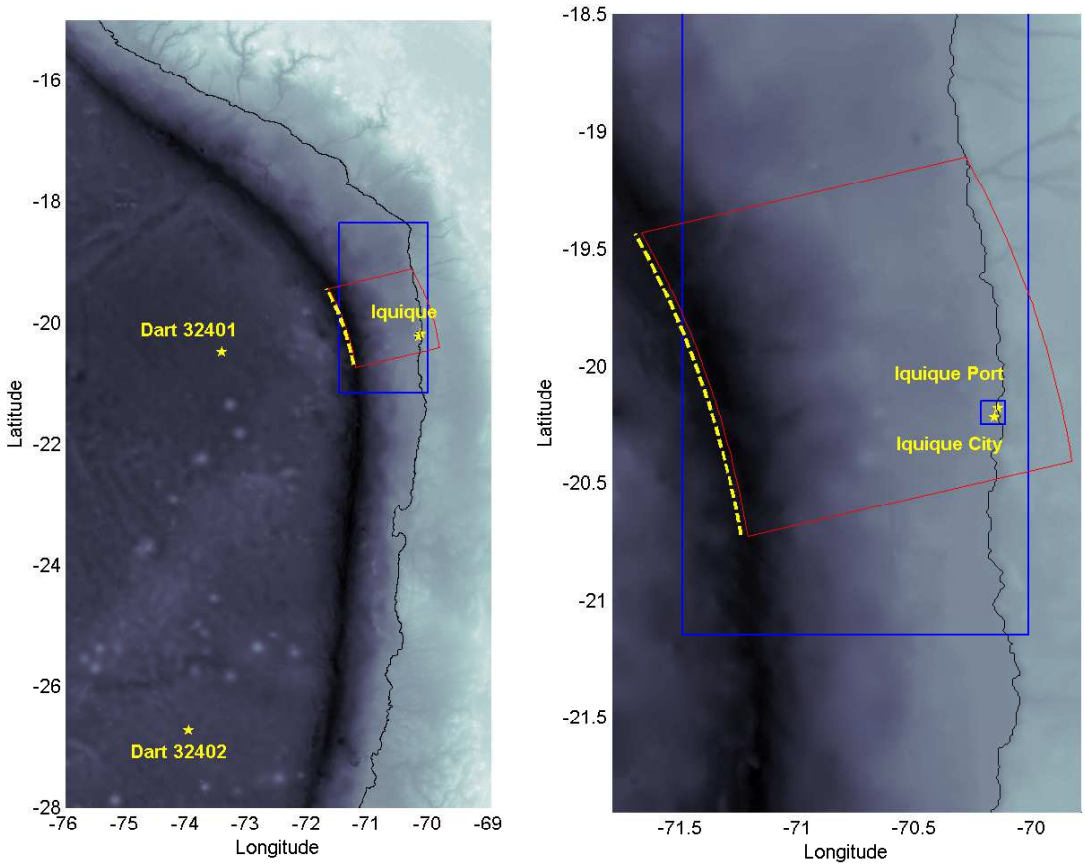


Figure 3.3: Problem setup of the 2014 illustration case. The region within the red lines corresponds to the projection of the seismogenic region on the horizontal plane, where hypothetical earthquakes with random distances to the trench are generated. The trench is indicated as the yellow segmented line. The left panel indicates the domain where tsunami wave propagation is resolved, while the blue rectangles depict regions of tsunami model refinement by means of nested grids. The yellow stars are the locations where tsunamis are assessed.

The slip standard deviation σ_s, on the other hand, is obtained from the scaling relation proposed by Raghukanth and Sangeetha (2016) in terms of the earthquake magnitude (bottom panel of Fig. 2.10). For the 2014 illustration case the Log-normal parameters have been estimated as $\bar{S}$ = 3.08 m and σ_s = 1.5 m.

The second property of the slip random field is the covariance function $c(\xi) = E[(S(x+\xi) - \bar{S})(S(x) - \bar{S})]$, where ξ is distance between two locations of the slip field. According to the analysis of Mai and Beroza (2002), in this thesis we adopt the Von Karman covariance function of Eq. 2.44. The correlation lengths are obtained with the scaling relations of Raghukanth and Sangeetha (2016) (upper panels of Fig. 2.10). The correlation lengths for the 2014 illustration case are a_x = 10.15 km (along dip direction) and a_y = 16.35 km (along strike direction). H_{VK} has been specified as 1.0, so that the high wave number components in the spectrum domain have a decay of k^{-2} (Mai and Beroza, 2002).

The centroid location of the rupture area $\vec{x}_c$ is defined as a random vector. Because of the lack of studies describing the likelihood of faults locations within a seismogenic region, we assume that all locations are equally likely to host an earthquake. This assumption corresponds to a random location with uniform distribution.

3.3 Sample generation

The generation of random vector samples for the earthquake location is straightforward. Toolboxes from many programming languages can be used to this purpose. For instance, the MATLAB function **mvnrnd** can be used to generate Gaussian random vectors. The generation of samples of a random slip field,

however, is rather complex.

Our objective is to develop a method for generating samples of the slip random field $S(x)$, $x \in \mathbb{R}^2$. The slip has finite variance and, therefore, can be modeled by means of a K-L expansion (Grigoriu, 2012). Since the samples of this field represent possible seismic events, they must be bounded, so that $S(x)$ cannot be modeled by a Gaussian field. We adopt a translation model to generate samples of $S(x)$. Thus, we first generate Gaussian field samples, $G(x)$, which are then transformed to samples of $S(x)$.

The 2014 application case considers a slip field with a Log-normal marginal distribution and a Von Karman covariance function. The Log-normal cdf, F, is given by,

$$F(\alpha) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{ln(\alpha) - \mu}{\sigma\sqrt{2}}\right)\right], \tag{3.1}$$

and its probability density function is given by,

$$f(\alpha) = \frac{1}{\alpha\sigma\sqrt{2\pi}} e^{-\frac{(ln(\alpha)-\mu)^2}{2\sigma^2}}. \tag{3.2}$$

The Gaussian cdf, on the other hand, is given by,

$$\Phi(\alpha) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{\alpha - \mu}{\sigma\sqrt{2}}\right)\right], \tag{3.3}$$

The parameters μ and σ are associated with the expected slip $\bar{S}$ and slip standard deviation σ_s of $S(x)$ by the expressions,

$$\mu = ln\left(\frac{\bar{S}}{\sqrt{1 + \frac{\sigma_s^2}{\bar{S}^2}}}\right) \; ; \; \sigma = \sqrt{ln\left(1 + \frac{\sigma_s^2}{\bar{S}^2}\right)}. \tag{3.4}$$

Eq. 3.1 and Eq. 3.3 can be then inserted in the translation model of Eq. 2.41. The covariance function of the Gaussian random field $c_G(\xi)$, on the other

hand, can be related analytically with the covariance function of the Log-normal random field $c(\xi)$ by,

$$c_G(\xi) = ln\left(1 + \frac{c(\xi)}{\bar{S}^2}\right). \tag{3.5}$$

It is important to remark that for a few distributions, such as the Log-normal of above, analytical expression can be obtained to relate the covariance functions. For most distributions, though, analytical expressions are not affordable. Moreover, for some target properties $F(\alpha)$ and $c(\xi)$, an exact solution for $c_G(\xi)$ does not exist. Grigoriu (1998) proposed some approximate approaches to relate $c(\xi)$ and $c_G(\xi)$ for any target distribution $F(\alpha)$.

3.3.1 Accuracy of sample generation.

The proposed method for earthquake sample generation has two sources of error, a spatial discretization error and a truncation error. They are related with the numerical solution of the eigenvalue problem in Eq. 2.39 and the truncation of the K-L expansion, respectively.

Spatial discretization: The eigenvalue problem is solved numerically by applying the Nystrom method (see Section 2.4.2). This method discretizes the random field into a grid with constant spacing. The integral equation problem of Eq. 2.39 is thus approximated by,

$$\sum_l \sum_m c_G(x_{lm} - x_{jk})\psi(x_{lm}) = \lambda\psi(x_{jk}), \tag{3.6}$$

where c_G adopts the form of a covariance matrix and x_{jk} and x_{lm} are locations in the discrete grid of R. The cells have dimensions Δx and Δy, perpendicular

and parallel to the trench direction, respectively. Note that the cells constitute sub-faults.

The error of the discretization depends on the frequency content of the samples of $S(x)$. This latter is represented by the power spectral density of the Von Karman covariance function (Goff and Jordan, 1988),

$$P(k_x, k_y) = \frac{4\pi\sigma_s^2 H_{vk} a_x a_y}{(1 + a_x^2 k_x^2 + a_y^2 k_y^2)^{1+H_{vk}}}, \tag{3.7}$$

where the total power is equivalent to the variance σ_S^2. Due to the Nyquist criterion, a discretized random field with sub-fault sizes $\Delta x \Delta y$ can only resolve wave numbers $|k_x| < \pi/\Delta x$ and $|k_y| < \pi/\Delta y$. This implies that a portion of the total power, associated with higher wave numbers, is lost in the discretization. The relative power loss is defined here as the discretization error, and is given by,

$$\epsilon_d = \frac{\int_{-\frac{\pi}{\Delta y}}^{\frac{\pi}{\Delta y}} \int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} P(k_x, k_y) dk_x dk_y - \sigma_s^2}{\sigma_s^2}. \tag{3.8}$$

Fig. 3.4 shows ϵ_d for different discretizations, normalized by the correlation lengths a_x and a_y.

Truncation error: The dominant K-L terms of the slip random field are associated with the top eigenvalues, which in turn are associated with greater length scales (small wave numbers). It is also noted that covariances with longer correlation lengths require fewer K-L terms for a given accuracy. Furthermore, if the samples are used as inputs of functions that filter high wave numbers, the number of relevant eigenmodes can be further reduced. This is the case of vertical

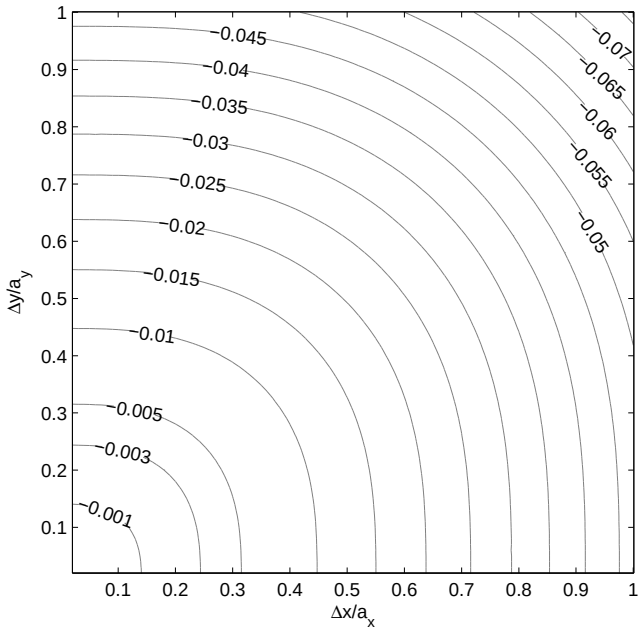


Figure 3.4: Discretization errors ϵ_d in terms of sub-fault size.

deformation solutions obtained with the Okada model, in which eigenmodes associated with small length scales do not affect the solutions. It has been also observed that this filtering effect is less effective for shallow earthquakes. For instance, Goda (2015) analyzed identical earthquakes with different fault depths. He noted that shallower earthquakes were associated with sharper seabed deformations patterns. This behavior is also observed in the comparison of Fig. 2.3. The truncation error can be defined as the loss of variance in the slip random field, as the same way as the discretization error. Alternatively, we can also define the truncation error as the loss of accuracy of the Okada solution when the K-L terms are reduced. The latter is adopted in this thesis.

LeVeque et al. (2016) analyzed the number of terms of the K-L expansion required to accurately model seabed deformation samples in the Cascadia Sub-

duction Zone. They discretized the rupture area by 865 sub-faults. They concluded that for a slip modeled with (1) a rupture area top boundary at 5 km, (2) an exponential covariance function and (3) a correlation length equal to 40% of the rupture area dimensions, 60 eigenmodes (K-L terms) are required to accurately compute samples. However, since the covariance functions, correlation lengths and rupture areas differ for each problem, the slip samples might require different discretizations (sub-faults) and number of K-L terms for a given accuracy. Hence, in this section we propose two criteria to discretize the rupture area and to select the K-L truncation, respectively.

First, we specify the sub-fault sizes for the random field discretization, based on the errors of Fig. 3.4. For the 2014 illustration case we use sub-fault dimensions $\Delta x = a_x/5$ and $\Delta y = a_y/5$, which produce an ϵ_d of 0.3%. Secondly, we measure the loss of accuracy in the Okada solution as a means to determine the minimum number of relevant K-L terms. Instead of examining every earthquake sample, we assess only the shallowest earthquake samples since they require more K-L terms. In the 2014 illustration case we selected the 25 shallowest earthquake samples (with their top boundaries closest to the depth of 1 km) and obtained their Okada solutions for the vertical seafloor deformation. Subsequently, we reduced the number of K-L terms for each of the 25 samples in order to observe the effects of the K-L expansion truncation. Fig. 3.5 shows the maximum error in seafloor deformation, normalized by the maximum (positive) seafloor deformation solved for each sample. The maximum error is defined as the maximum seafloor displacement difference between two estimations: one is based on the specified truncated K-L expansion and the other is based on the K-L expansion using all the eigenmodes computed in Eq. 3.6, which in this case is 1575 (using a discretization $\Delta x = a_x/5$ and $\Delta y = a_y/5$). We observe that keep-

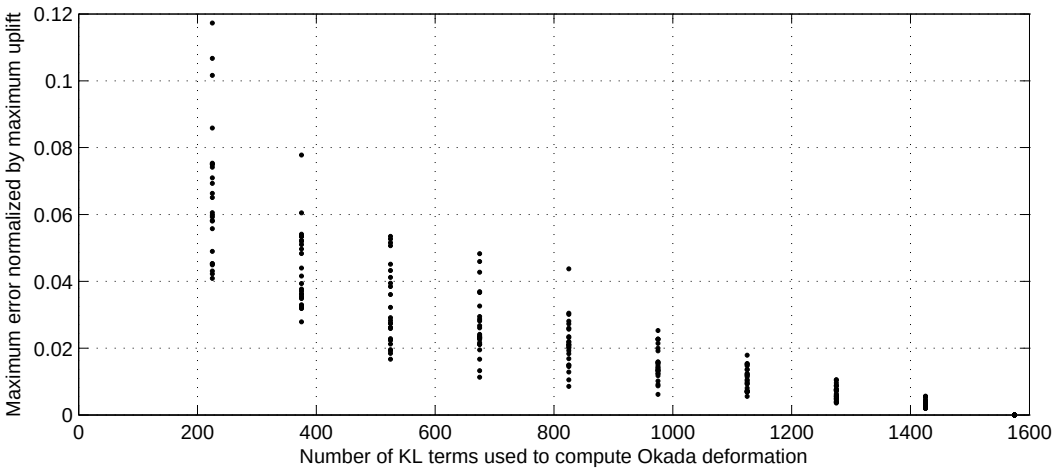


Figure 3.5: Errors in the vertical seafloor deformation for different number of K-L terms in the 2014 illustration case. The rupture area has been discretized into 1575 sub-faults with size $\Delta x = a_x/5$ by $\Delta y = a_y/5$. 25 samples are analyzed for each truncation.

ing 600 K-L terms gives a representation of Okada model solutions, with errors smaller than 5%. Fig. 3.6 shows the first ten eigenmodes ψ corresponding to the ten largest eigenvalues. The eigenmodes are normalized so they have unit ℓ^2-norm.

It is important to reiterate here that the sub-fault discretization and the truncation of the K-L expansion are specific for the illustration case and the analysis should be repeated for each new case of study. Another noteworthy aspect of our illustration case is that we have assumed that the water surface mimics the vertical seafloor deformation during an instantaneous fault slip. This assumption does not hold for seafloor deformation patterns whose wavelengths are comparable to the water depth. In this illustration case, we verified that dominant wavelengths of the seafloor deformation samples are more than three times the water depth (see Fig. 2.3), which is a reasonable length to adopt the men-

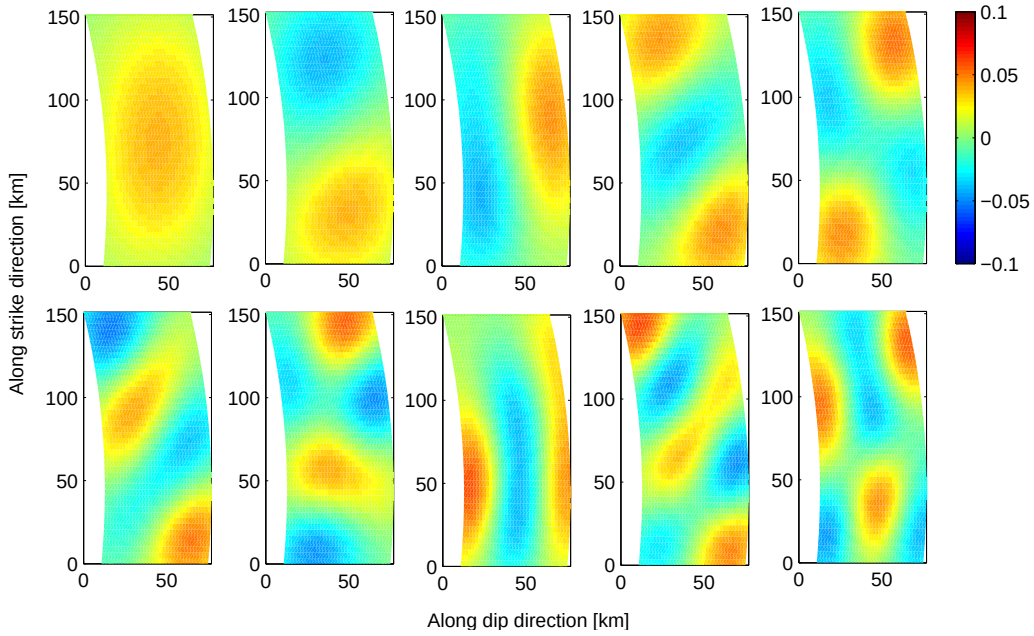


Figure 3.6: First 10 eigenmodes ψ of the K-L expansion, ordered from left to right and top to bottom. These eigenmodes are associated with the Gaussian random field and non-rectangular rupture area used in the 2014 illustration case. The eigenmodes are composed by sub-faults with size $\Delta x = a_x/5$ by $\Delta y = a_y/5$ and they are normalized so they have unit ℓ^2-norm.

tioned assumption (Kajiura, 1963). However, as stated by Geist and Dmowska (1999), special attention should be paid to very shallow rupture areas. In such cases, vertical seafloor deformations may contain significant short wavelength patterns which are not transmitted to the water surface. It is important to remark that Kajiura (1963) proposes some procedures to filter such a short wavelength patters to build reasonable tsunami initial conditions.

3.3.2 Analysis of the generated samples.

For the 2014 illustration case we have generated 10,000 earthquake samples to assess their statistical properties. Fig.3.7 shows 12 earthquake samples. The left panel of Fig.3.8 shows the histograms of the slip of each sub-fault as grey curves and the target Log-normal probability density function $f(x)$ (Eq.3.2) as a dashed black curve. The average root mean square error of the histograms with respect to $f(x)$ among the sub-faults is $\epsilon_{hist} = 5 \times 10^{-3}$. The average errors of the mean, standard deviation and skewness among sub-faults are 0.1%, 0.6% and 2.0%, respectively, relative to the moments of the target Log-normal distribution. We also check the consistency of the covariance function among the samples. Fig. 3.9 shows the covariance normalized by the slip variance for the 10,000 samples. The root mean square error of the covariances, with respect to the Von Karman function, is $\epsilon_c = 0.024\ m^2$. This corresponds to 1.1% of the maximum covariance.

We have demonstrated that the present methodology generates samples which are consistent with the target probability properties. To weigh its pertinence, we analyze the severity of the inconsistencies identified in existing methodologies. First, we scale the slip samples to preserve the same earthquake magnitude. We use the 10,000 samples of the 2014 illustration case. As we observe in Fig. 3.10(a), the errors in the histrograms, ϵ_{hist}, increase to more than the doubles. The errors in the covariance values, on the other hand, increase by one order of magnitude. As a second experiment, we assess the impact of generating slip samples by using a rectangular rupture area which is then deformed to the non-rectangular rupture area of the 2014 illustration case. This procedure is commonly required in Fourier-transform-based methods to gener-

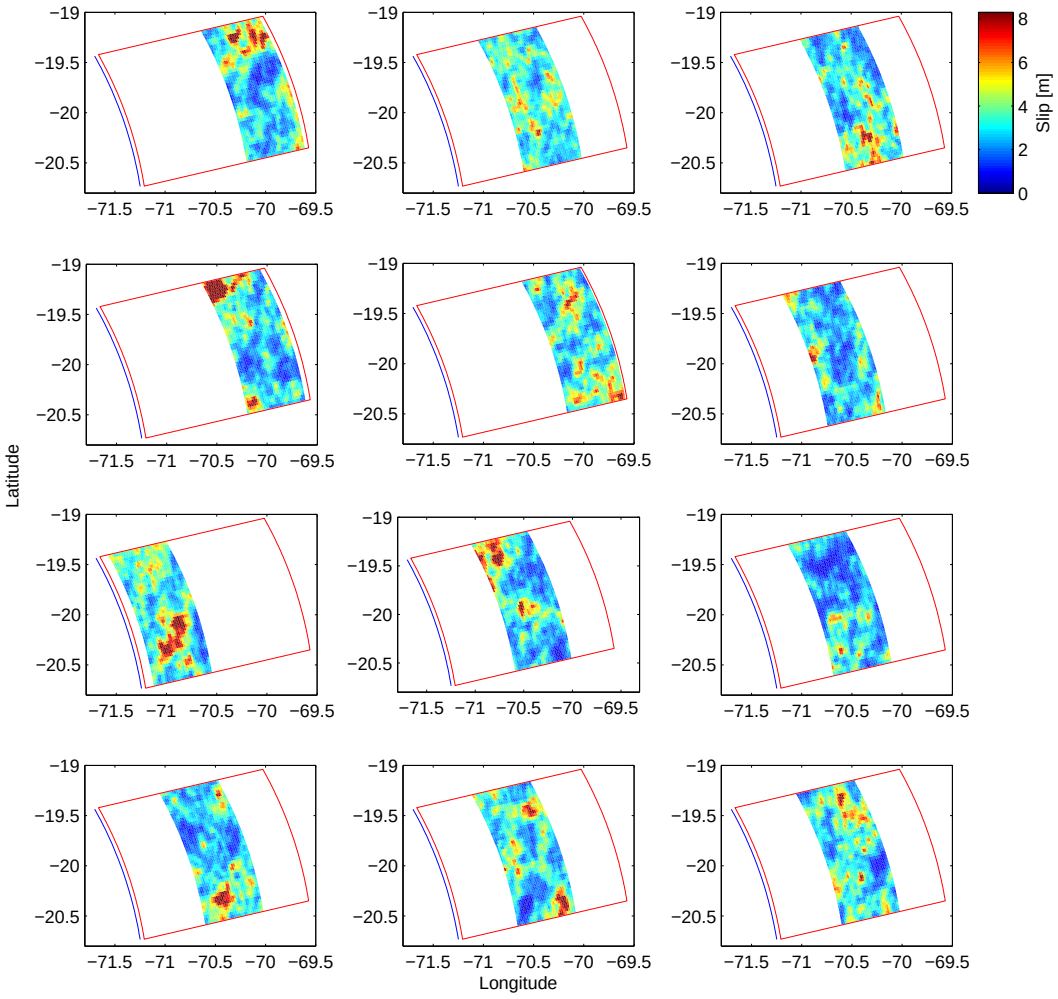


Figure 3.7: Projection of 12 samples of the random slip distribution S on the horizontal plane, with random distance to the trench for the 2014 illustration case. The expected earthquake magnitude of the random field is M_W 8.0. The dimensions of the rupture area are obtained from scaling relations of Blaser et al. (2010). The marginal distribution is Log-normal with $\bar{S}$ = 3.08 m and σ_s = 1.50 m. The covariance function is Von Karman with a_x = 10.15 km and a_y = 16.35 km, obtained from scaling relations (Raghukanth and Sangeetha, 2016). The red and blue lines indicate the boundaries of the seismogenic region and the trench, respectively.

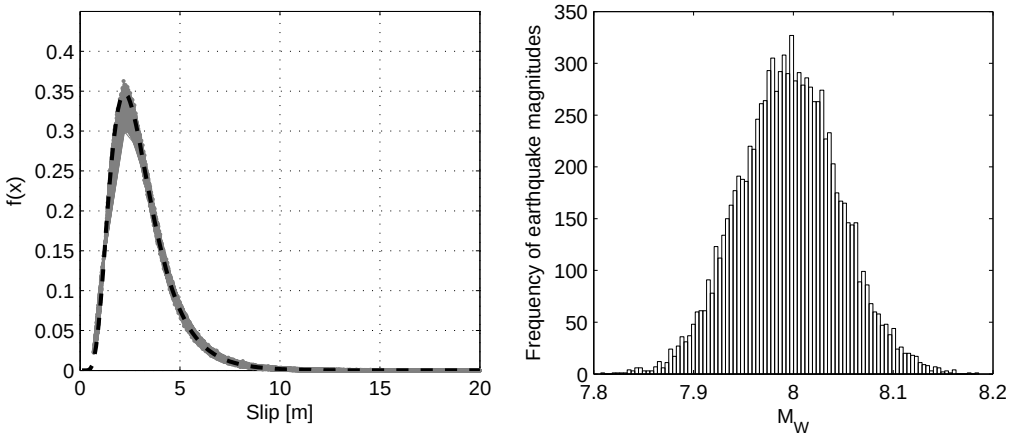


Figure 3.8: Left: Histograms of the slip for each sub-fault as grey curves and target Log-normal distribution as a dashed black curve. Right: Histogram of the earthquake magnitude of samples. The probability properties has been obtained with 10,000 samples.

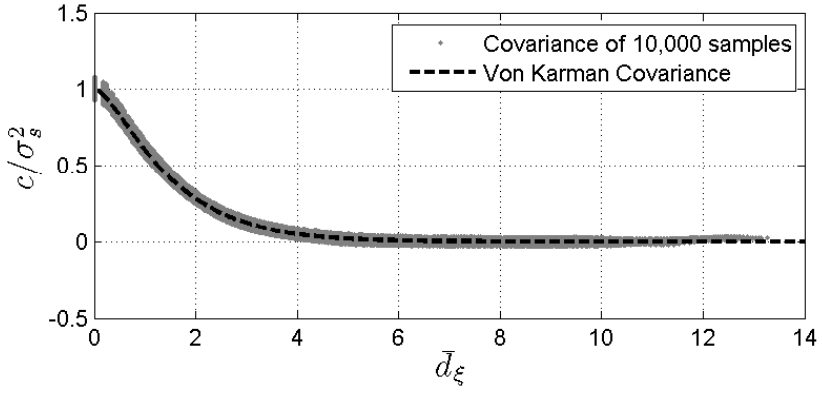


Figure 3.9: Target Von Karman covariance function and terms of the covariance matrix obtained with 10,000 samples. The parameter $\bar{d}_\xi$ is defined in Eq. 2.44.

ate samples of non rectangular rupture areas (e.g. Li et al. (2016)). As we observe in Fig. 3.10(b), the histograms have the same errors as the present methodology. However, the error in the covariance increases to more than the doubles. It is important to note that the rupture area of the 2014 illustration case is not significantly curved and in more complex rupture areas we expect larger errors in Fourier-transform-based methods. Finally, as a third experiment, we adopt the procedure suggested by LeVeque et al. (2016) to reduce the earthquake magnitude variability. They noted that some K-L modes have large non-zero mean over the rupture area and, thus, they contribute to the earthquake magnitude variability of samples. It has been observed that the first K-L mode (defined as zero-th mode in LeVeque et al. (2016)) has a major contribution to this variability. Hence, the earthquake magnitude variability can be significantly reduced by removing this latter term from the K-L expansion. Fig. 3.10(c) shows the probability properties of samples generated without the first K-L term. We observe that the errors are as large as those obtained with the scaling procedure. Conclusively, these experiments demonstrate that the three inconsistencies identified in existing methodologies are relevant. Our methodology, therefore, significantly improves the slip sampling generation, with respect to the target probability properties.

In the right panel of Fig. 3.8 we show that the sample earthquake magnitudes are consistently centered towards the expected magnitude of the defined random field, i.e., M_W 8.0 (average with an error of 0.08% for 10,000 samples). However, the individual earthquake samples have different magnitudes. The explanation for the differences is given as follows. The statistics of slip over the rupture area for a given sample (i.e. histogram of sub-fault slips) differs from the marginal probability distribution when the random field is not ergodic

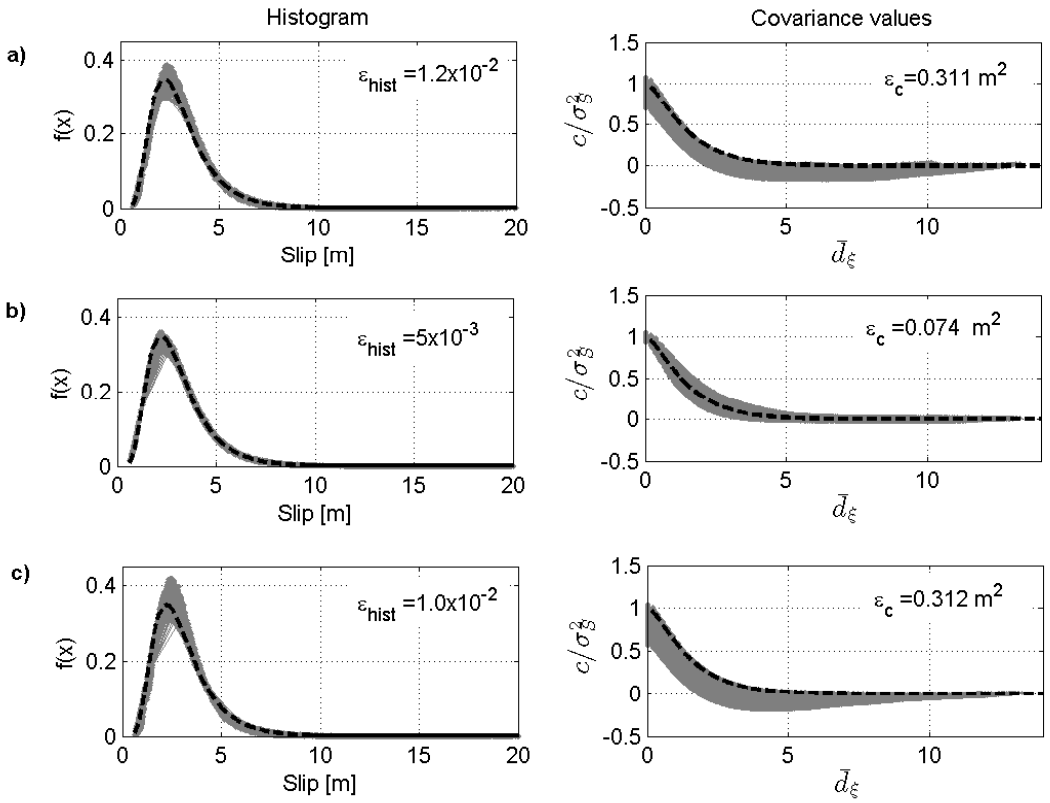


Figure 3.10: Probability properties of samples generated by using common procedures of existing methodologies. Grey curves are probability properties of generated samples and black dashed curves are target probability properties. (**a**) Scaling of samples to preserve earthquake magnitude. (**b**) Samples generated by using a rectangular random field which is then deformed to a non-rectangular rupture area. (**c**) Samples generated without the first mode of the K-L expansion with the purpose to reduce earthquake magnitude variability. The probability properties of the samples generated with the present methodology are shown in left panel of Figure 8 and Figure 9. The errors of the present methodology are $\epsilon_{hist} = 5 \times 10^{-3}$ and ϵ_c=0.024 m^2.

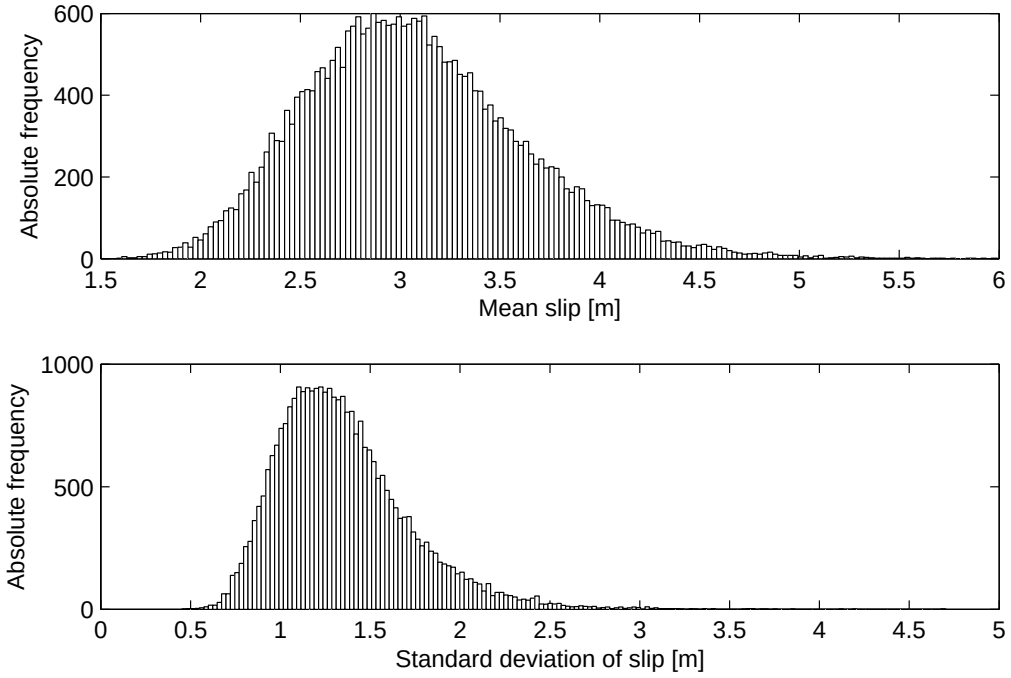


Figure 3.11: Histogram of slip average and standard deviation over the rupture area for slip samples of the 2014 illustration case.

or when characteristic correlation lengths are comparable to the dimensions of the rupture area. The latter is the case of the slip field. Thus, spatial averages and spatial variances of the slip over the sub-faults differ from the mean slip and variance of the marginal distribution, respectively, as we observe in the histograms in Fig. 3.11. The sample spatial slip average, in particular, is associated with the seismic moment, which in turn, is related to the earthquake magnitude.

Since the earthquake magnitude of the slip samples differs from the magnitude used in the scaling relations, those samples are inconsistent. In other words, these slip samples are related to different rupture area dimensions, correlation lengths and standard deviations in the scaling relations. Here, we check the severity of this inconsistency by comparing the dispersion of the sample

magnitudes with the scattering of the data used to obtain the scaling relations. In Fig. 3.12, the grey "×" points correspond to the fault geometries of past earthquakes collected by Blaser et al. (2010), which are used to find the scaling relations of the rupture area dimensions. The black points in the same figures represent the rupture area length and width of samples (specified by the scaling relation and the expected earthquake magnitude) versus the actual sample magnitudes M_W. The sample magnitudes do not deviate significantly from the fitted curve as compared with data scattering. Hence, we conclude that the inconsistency in the use of the scaling relation for the dimensions of the rupture area is not severe. This conclusion is also extended for the correlation lengths and slip variance. For these latter parameters, though, some relevant aspects have been observed. The scaling relations of Raghukanth and Sangeetha (2016) for the correlation lengths and slip standard deviation have been determined from data describing past earthquake events. In one hand, the correlation lengths of the past earthquake events are obtained by applying a Fourier transform into the event slip distribution and by then fitting the curve to a Von Karman function with calibrated correlation lengths. We remark that the resulting spectra, and correlation lengths, are similar but not equivalent to the spectra and correlation lengths (a_x and a_y) of the actual covariance function of the slip field. Unlike the inferred parameters, a_x and a_y are the result of an average behavior among more than one sample. It is also important to mention that the procedure to determine the sample correlation lengths is challenging because rupture dimensions are comparable to the correlation lengths. As a consequence, the Fourier spectrum is not well described due to spectral leakage originated by the finite length of the rupture area along the strike and dip directions (also known as windowing effect). We observed that acceptable noise-free

slip spectra were only obtained for earthquake magnitudes M_W greater than 7. It is also observed that Mai and Beroza (2002) and Raghukanth and Sangeetha (2016) have applied interpolations between slip distribution sub-faults with the purpose to increase the Nyquist frequency (maximum resolvable frequency in spectra), or have padded zeros to refine the frequency resolution of the slip spectra (i.e. refine the spectra). These procedures, however, do not provide better quality or further information, as shown by the blue curves of Fig. 2.9. Hence, the existing approach for the determination of the correlation lengths from past earthquakes might lead to errors. The scaling relation of the slip standard deviation, on the other hand, has been obtained by adopting the standard deviation of the slip over the rupture area (i.e. spatial slip standard deviation), which is also similar but not equivalent to the slip standard deviation σ_S of the marginal distribution. This latter is actually the slip standard deviation at a specific spatial location. Fig. 3.13 and Fig. 3.14 show the correlation lengths and standard deviation of 100 simulated samples with different expected earthquake magnitude (black point). Note that the samples have varying sample correlation lengths and varying spatial slip standard deviation, as we explain above. Despite all the mentioned inconsistencies, the magnitudes of the simulated earthquake samples do not deviate significantly from the fitting curves as compared with the earthquake data dispersion (grey "×" points).

3.4 Uncertainty propagation

In this section we describe the methodology to propagate the earthquake uncertainty into tsunami simulation results. We use the Stochastic Reduced Order Model (SROM), proposed by Grigoriu (2009) and described in Section 2.4.3.

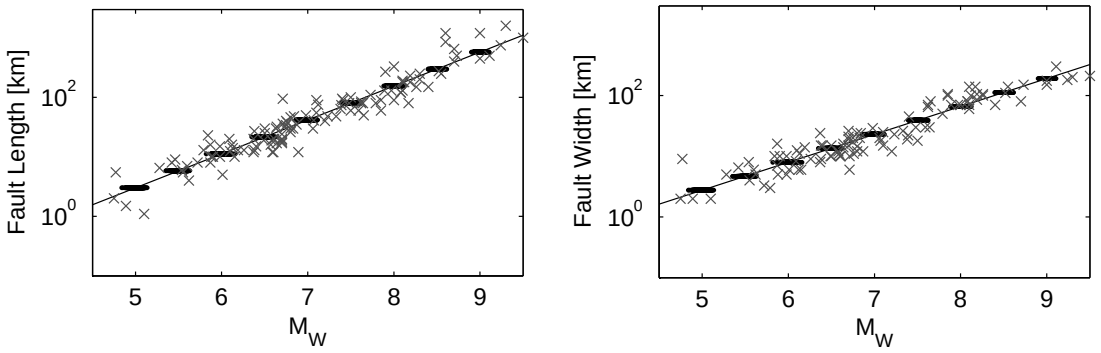


Figure 3.12: Black points: Fault length and width of samples in terms of sample magnitude, obtained by using the scaling relations of Blaser et al. (2010) and by the specification of an expected magnitude M_W equals to 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5 and 9. For each expected magnitude 1,000 samples are obtained. Note that the sets of 1,000 points are so close that they appear to be a black thick line. The grey "×" points are earthquake data used in Blaser et al. (2010). The black lines are the fitted curves.

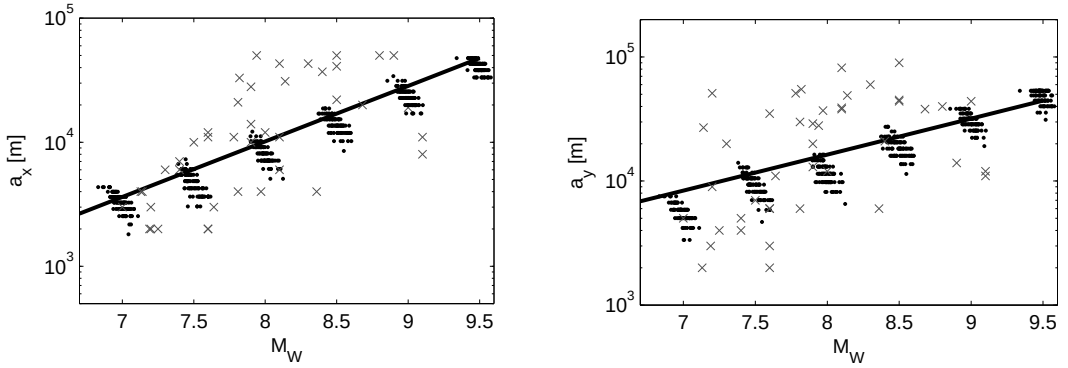


Figure 3.13: Black points: Correlation lengths of simulated samples in terms of the sample magnitude, obtained by using a Fourier transform. The black fitted lines correspond to the scaling relations of a_x and a_y, provided by Raghukanth and Sangeetha (2016). The grey "×" points are earthquake data.

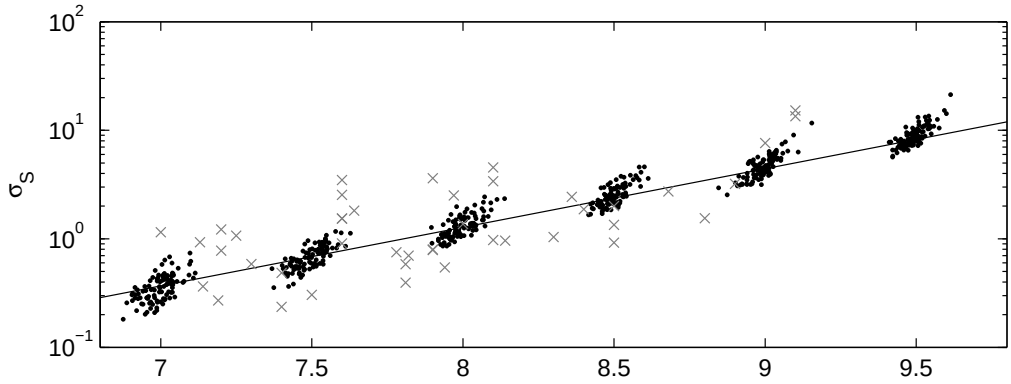


Figure 3.14: Black points: Standard deviation of slip over the rupture area σ_{sample} of samples in terms of sample magnitude. The earthquake samples were obtained using σ_s of the scaling relations of Raghukanth and Sangeetha (2016) (black fitted line). The specified expected magnitudes M_W are 7, 7.5, 8, 8.5, 9 and 9.5. For each expected magnitude we ran 100 samples. σ_s is in meters. The grey "×" points are earthquake data.

SROM can be seen as an optimized version of a classic Monte Carlo simulation, which estimates more accurate statistics of model outputs for a given number of samples.

3.4.1 SROM of tsunami model input

As described in Section 2.4.3, SROM solves the optimization problem of Eq. 2.42, in which the discrepancies between target probability properties of inputs and the statistics of a set of m samples is minimized. The objective function considers relevant probability properties for the problem which is analyzed. Since the initial conditions for tsunami simulation are the prescribed water surface displacement approximately mimicking the seafloor deformation, we use the probability properties of this latter in the objective function. Note that these lat-

ter probability properties also depend on $\tilde{Z}$ and $\vec{p}$ (i.e. earthquake random characteristics). The target probability properties of the seafloor deformation are obtained by solving the Okada model for each of the 10,000 earthquake samples generated in the previous section. Then, the discrepancies - between the target and the statistics of m initial conditions - in the objective function are given by,

$$
\begin{aligned}
e_1(\tilde{Z},\vec{p}) &= \sum_i \int_{I_i} (\tilde{F}_\eta(x_i) - F_\eta(x_i))^2 dx_i \\
e_2(\tilde{Z},\vec{p}) &= \sum_i \sum_{q=1}^{4} \left(\tilde{\mu}^q_{\eta,i} - \mu^q_{\eta,i}\right)^2 \\
e_3(\tilde{Z},\vec{p}) &= \sum_{i,j;j>i} \left(\tilde{c}_{\eta,i,j} - c_{\eta,i,j}\right)^2
\end{aligned}
\tag{3.9}
$$

where I_i is the range of possible values of the seafloor deformation at the $i-th$ location, $F_\eta(x_i)$ and $\tilde{F}_\eta(x_i)$ are the cdf's of the seafloor deformation at the $i-th$ location associated with Z and $\tilde{Z}$, respectively. $\mu^q_{\eta,i}$ and $\tilde{\mu}^q_{\eta,i}$ are the mean (q=1), variance (q=2), skewness (q=3) and kurtosis (q=4) of the seafloor deformation at the $i-th$ location associated with Z and $\tilde{Z}$, respectively. $c_{\eta,i,j}$ and $\tilde{c}_{\eta,i,j}$ are the covariance of the seafloor deformation between the i-th and j-th location associated with Z and $\tilde{Z}$, respectively. The values for the weights α_1, α_2 and α_3 in Eq. 2.43 can be obtained by trial and error so that the contributions of e_1, e_2 and e_3 in e_t have the same order of magnitude. In the 2014 illustration case we use $\alpha_1 = 1.0$, $\alpha_2 = 0.2$ and $\alpha_3 = 0.3$. We observe that these values vary for each new case of study.

The 2014 illustration case adopts the sub-optimal procedure of Grigoriu (2012), with $n = 10,000$, $m = 100$ and $n_{set} = 100$. Fig. 3.15 shows the minimum discrepancies in initial condition statistics in terms of n_{set}. The discrepancies have been normalized by the discrepancy of the first optimized set of samples. Note that a considerable discrepancy reduction is achieved for n_{set} <50. This

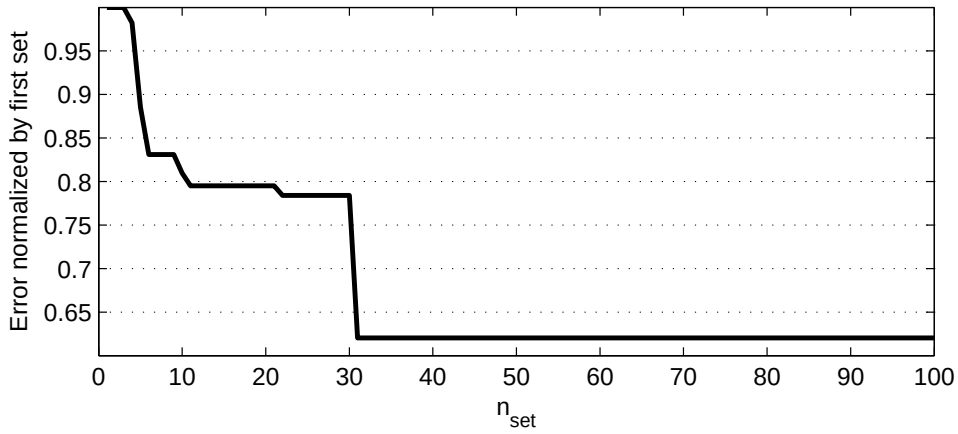


Figure 3.15: Discrepancies between the $m = 100$ sample statistics and the target probability properties in terms of n_{set} of the sub-optimal procedure, used to find the optimized SROM. Discrepancies are normalized by the discrepancy of the first tested set.

procedure was repeated 10 times, yielding similar behavior. In this illustration the optimization problem took 12 hours using a conventional I7 processor computer. It is important to remark again that discrepancies in Eq. 4.6 decrease for larger values of m and n_{set}. Therefore the user must select the most efficient size of the SROM and perform sensitivity analyses to define m and n_{set}. On the other hand, the objective function can be also modified to improve the convergence of the statistics of the tsunami model outputs.

3.4.2 SROM of tsunami model output

The SROM $\tilde{Z}$ is then used to construct a SROM $\tilde{Y}$, which is the stochastic model output. Let $Y = h(Z)$ be the output of a stochastic tsunami model (e.g., random surface elevation, velocities, runup, etc.), with earthquake inputs Z. The

approximate solution $\tilde{Y}$ is given by the pairs:

$$\tilde{Y} = (\{\tilde{Y}_1, p_1\}, \{\tilde{Y}_2, p_2\}...\{\tilde{Y}_m, p_m\}) \tag{3.10}$$

where $\tilde{Y}_k = h(\tilde{Z}_k)$ are the solutions of a tsunami deterministic model using $\tilde{Z}_k$ as input, with same probabilities. Grigoriu (2012) showed that the errors of $\tilde{Y}$ approximating Y are bounded and related to the discrepancies between Z and $\tilde{Z}$. That means that $\tilde{Y}$ converges to the exact solution Y as the discrepancy between Z and $\tilde{Z}$ vanishes. $\tilde{Y}$ can be used to estimate distribution, moments and covariances of Y. Note that the SROM requires m runs of a deterministic tsunami model.

We used the m(=100) samples of the seafloor deformation of the 2014 illustration case as initial conditions for tsunami simulations. Furthermore, the bathymetry has been updated to account for the co-seismic deformation, which can be relevant in coastal areas. The tsunami model COMCOT was used. The simulation domain is shown in the left panel of Fig. 3.3 and has a grid resolution of 1.5 arcmin. Inside this grid, two nested grids are inserted enhancing the resolution to 0.3 arcmin (blue rectangle in left panel of Fig. 3.3) and 0.06 arcmin (blue square in right panel). This latter is equivalent to a 100 m resolution and is located in the coastal area of Iquique, where time histories of tsunami elevation will be obtained from the simulations. The model simulates bottom friction by adopting a quadratic law model with a constant Manning's number of 0.03. The time step used in the computation is 1 second. The $m = 100$ simulations were ran in 15 hours using a conventional I7 processor computer. In addition to the optimization procedure, the methodology required a total of 27 hours for the 2014 illustration case.

Fig. 3.16 shows the time histories of the simulated tsunami amplitudes

at four locations around the seismogenic region: two in front of the port (Lon-70.148^o Lat-20.204^o) and city of Iquique (Lon-70.170^o Lat-20.220^o) and two at offshore buoys locations (DART32401 at Lon-73.429^o Lat-20.473^o and DART32402 at Lon-73.983^o Lat-26.743^o), which are depicted in Fig. 3.3. The time histories of the earthquake samples are depicted as black curves. By using the probabilities obtained for the SROM $\tilde{Z}$, we have also computed the 50%, 15% and 5% exceedance probabilities of the time histories. These curves inform about the probability to exceed a tsunami elevation at a given time.

3.4.3 Analysis of the uncertainty propagation.

Our results will be first compared with the results of existing methodologies. Then, they are compared with the records of the 2014 tsunami.

Based on our numerical results, the maximum tsunami amplitudes for the $m = 100$ tsunami samples have been computed and used to elaborate probability exceedance curves of the maximum tsunami amplitudes. These curves are represented by the black lines in Fig. 3.17. In addition, we have indicated the value of coefficient of variation (CoV) of the maximum tsunami amplitude at each location, which is defined as the standard deviation divided by the mean, and can be interpreted as a metric of the uncertainty. To further evaluate the relevance of the probability properties of the slip, the proposed earthquake sampling method and the proposed uncertainty propagation method on final tsunami results, we perform three comparison analyses. First, we repeat the methodology presented above, but initially define an exponential target marginal distribution instead of a Log-normal marginal distribution. As we mentioned in

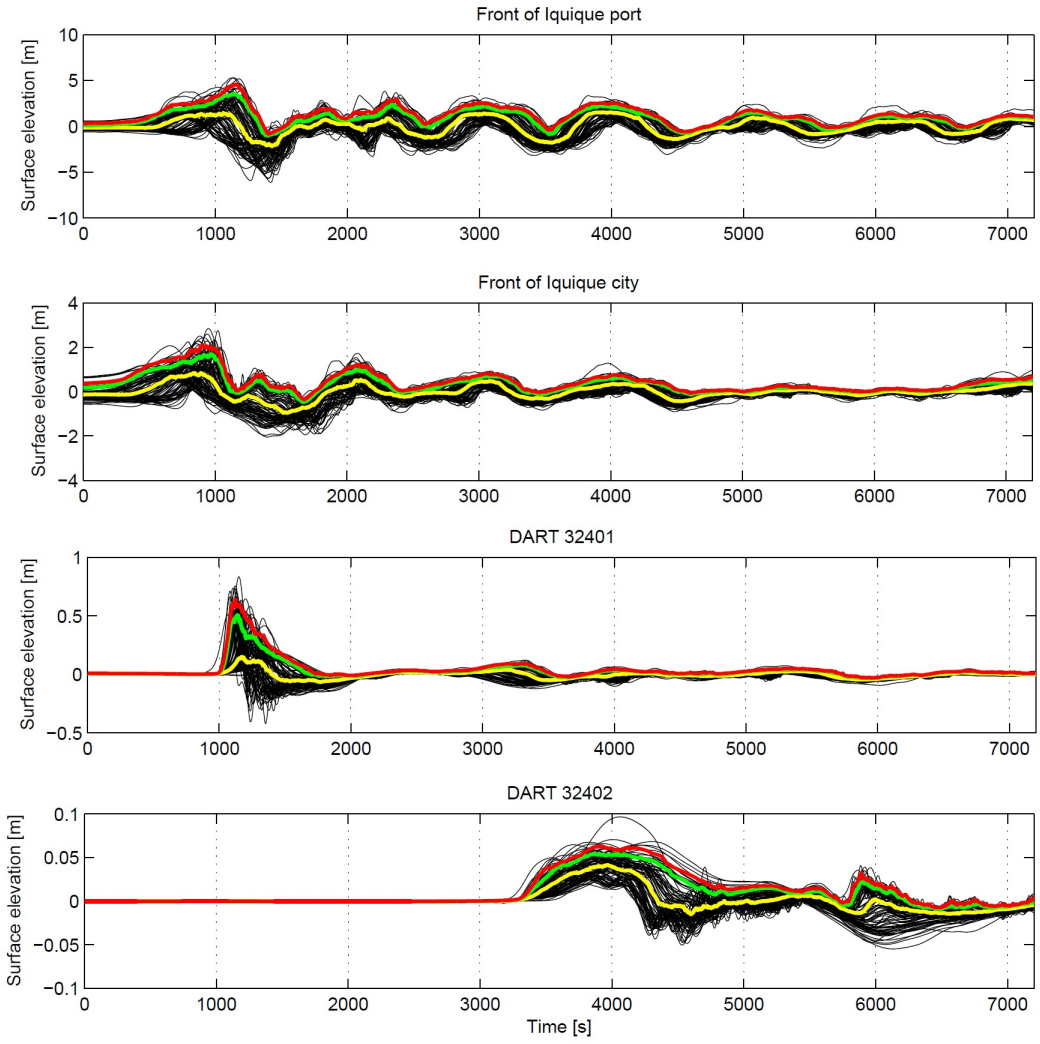


Figure 3.16: Tsunami wave elevations at 4 locations close to the rupture area. Black time histories corresponds to the m =100 tsunamis samples. The red, green and yellow curves indicated the 5%, 15%, 50% exceedance probabilities, respectively.

section 2.5.1, this is the most controversial probability property to define the slip distribution. The exponential distribution is observed in the spatial slip histogram of the 2011 Japan earthquake (see Fig. 2.8) and it is similar to the truncated exponential distribution proposed by Thingbaijam and Mai (2016). We remark that we verified that target Von Karman covariance function of the final earthquake samples is not modified as a consequence of the application of the translation model. The grey curves in Fig. 3.17 correspond to tsunamis generated with earthquakes whose slip follows an exponential marginal distribution. Note that the grey curves are associated with larger tsunami responses for small exceedance probabilities, as compared with the black curves (i.e. Log-normal marginal distribution). This comparison demonstrates that the adopted target marginal distribution for the slip has a significant effect on tsunami responses. Second, we adopt the post-sampling scaling procedure to preserve earthquake magnitude. We use this procedure because it is the post-sampling procedure most used in existing methodologies and because it shows the largest inconsistencies with probability properties, as shown in Fig. 3.10. Once the scaled samples are generated, the statistical properties of 10,000 samples are obtained, which are used in a new SROM. The orange solid curves in Fig. 3.17 are the exceedance probabilities associated with the scaled earthquake samples. We observe that the orange and black exceedance curves (non-scaled samples) differ for probabilities smaller than 10%. In general, the present method (black curves) shows larger probabilities for a given maximum tsunami amplitude. The differences are more relevant at offshore locations. This comparison demonstrates the importance of the method adopted to generate earthquake samples on tsunami hazard assessments. Third, to assess the performance and accuracy of the SROM, we have conducted one Monte Carlo simulation with 1,000 sam-

ples, which required 150 hours, and 40 Monte Carlo simulations with 100 samples, each one requiring 15 hours. Results are shown as the dashed red curve and green curves in Fig. 3.17, respectively. The SROM result with $m = 100$ is close to the result of the Monte Carlo simulation with 1,000 samples. The Monte Carlo results using 100 samples, conversely, have a significant scattering around the red and black curves, which are associated with errors of about 20% in the estimation of the maximum tsunami amplitudes with 10% exceedance probability. Note also that the SROM for this illustration case requires less than a fifth of the time of a classic Monte Carlo simulation (with 1,000 samples) to achieve similar accuracies. This comparison provides strong evidence that SROM is much more accurate than classic Monte Carlo simulations with the same number of tsunami simulations.

It is important to note that the 2014 event can be considered as a sample of the random earthquakes in the seismigenic region, and therefore, the field data should be contained inside the cloud of sample solutions. Thus, the tsunami measurements of the 2014 Iquique Earthquake provide a check, in some degree, for the methodology proposed in this paper. In Fig. 3.18 we observe that the SROM samples indeed contain the measurements of the leading and first waves of the tsunami in the assessed locations. The amplitude of the trailing waves are also well captured by the samples of our model, but with some phase mismatches. The phase mismatches are associated with sources of uncertainty which are not considered in this methodology. Some possible sources could be errors in the specification of some probability properties of the earthquake random slip and location or inaccuracies of the tsunami model due to the assumptions and simplifications adopted on it. Other inputs of the tsunami models can be also relevant sources of uncertainty. The bathymetry, in particular, is thought

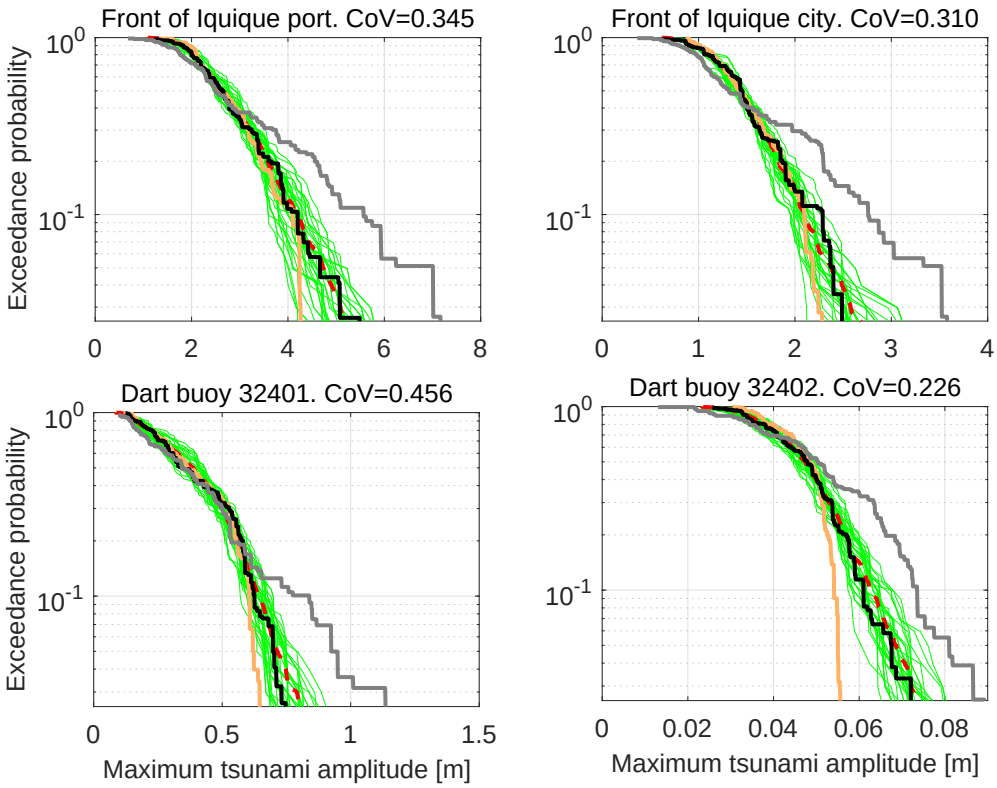


Figure 3.17: Exceedance probability curves of the maximum tsunami amplitude relative to the still water level at the four assessed locations. The black solid curve corresponds to SROM with m=100 earthquake samples, selected by the present method. The solid grey curves correspond to an SROM with m=100 earthquake samples, which are defined with an exponential marginal distribution instead of a Log-normal marginal distribution. The orange solid curves correspond to an SROM with m=100 earthquake samples, which are scaled to preserve earthquake magnitude. The green curves corresponds to 40 realizations of Monte Carlo simulations with 100 samples. The red dashed line corresponds to a Monte Carlo simulation with 1,000 samples and is viewed as reference of the true exceedance probability.

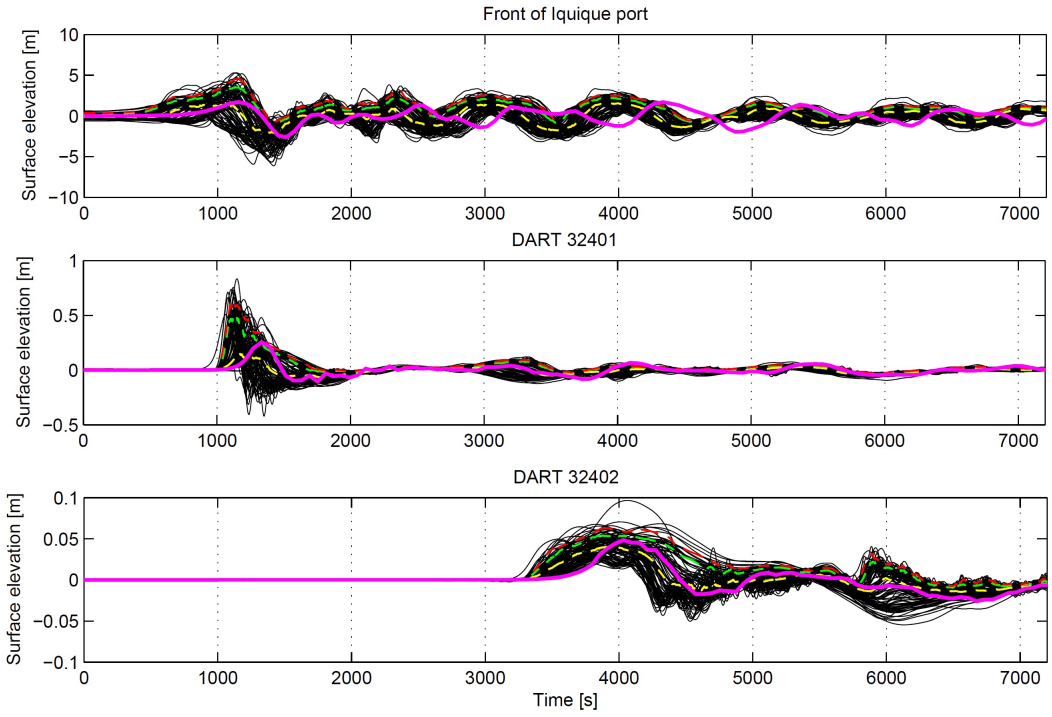


Figure 3.18: Comparison of the 100 SROM samples as black curves and measurements of the tsunami in 2014 as magenta curves. The red, green and yellow segemented curves correspond to the exceedance curves of Fig. 3.16.

to be a potential source of uncertainty and is analyzed in the next chapter.

CHAPTER 4

UNCERTAIN BATHYMETRY MODEL

This chapter presents a new method for the quantification of uncertainty in tsunami hazard assessments due to uncertain bathymetry. We first illustrate the relevance of the uncertain bathymetry by means of a motivational experiment. Then we present the new methodology. As the case of Chapter 3, we use the 2014 illustration case of tsunamis in North Chile to explain the methodology. We consider an uncertain bathymetry within mesh 2 and mesh 3, which have the finest resolutions of 0.3 arcmin (500 m) and 0.06 arcmin (100 m), respectively (blue rectangles in the left and right panels of Fig. 3.3). The uncertainties of mesh 1 are assumed to be zero because most of this mesh covers water depths of $\sim$ 3 km and the maximum standard deviation of the bathymetry uncertainty is $\sim$ 0.7 km (about 1/4). Under these conditions, bathymetry uncertainties are comparatively small and do not contribute to tsunami uncertainties significantly, as it is observed in the motivational experiment below.

4.1 Motivation

In this motivation we assess the effect of uncertain bathymetry on a simplified tsunami propagation problem. Let us consider a constant depth ocean with depth h_1 and a linear sinusoidal wave with unit amplitude and wavelength L_W propagating through shallow waters. Let us also consider a section in the ocean that has not been surveyed. This unsurveyed section has width W_r and unknown constant depth h_2, as seen in Fig. 4.1. The uncertain bathymetry section might be a ridge ($h_2 < h_1$) or a trench ($h_2 > h_1$).

Mei et al. (2005) presents the solution for long and linear sinusoidal waves after to cross a finite width ridge (or trench). According to Mei et al. (2005), a wave normally propagating to the ridge (or trench) has the following solution before (η_1), during (η_2) and after (η_3) to cross the ridge (Fig. 4.1),

$$\begin{aligned} \eta_1 &= \left(e^{ik_1(x+W_r/2)} + R_r e^{-ik_1(x+W_r/2)}\right) e^{-i\omega_i t}, \\ \eta_2 &= \left(Ae^{ik_2 x} + Be^{-ik_2 x}\right) e^{-i\omega_i t}, \\ \eta_3 &= \left(Te^{ik_3(x-W_r/2)}\right) e^{-i\omega_i t}, \end{aligned} \tag{4.1}$$

where ω_i is the wave angular frequency, R_r is the amplitude of the wave reflected in the ridge (trench), A and B are the amplitudes of the waves propagating forward and backward over the ridge (trench), respectively, and T is the amplitude of the transmitted wave. Note that A,B, R_r and T might be complex numbers. k_1, k_2 and k_3 are the wavenumbers before, during and after to cross the ridge (trench), respectively, and $k_1 = 2\pi/L_W$. Mei et al. (2005) specified the following boundary conditions for the problem,

$$\begin{aligned} \eta_{1(x=-W_r/2)} = \eta_{2(x=-W_r/2)};\ \eta_{2(x=W_r/2)} = \eta_{3(x=W_r/2)}, \\ h_1 \frac{\partial \eta_{1(x=-W_r/2)}}{\partial x} = h_2 \frac{\partial \eta_{2(x=-W_r/2)}}{\partial x};\ h_2 \frac{\partial \eta_{2(x=W_r/2)}}{\partial x} = h_1 \frac{\partial \eta_{3(x=W_r/2)}}{\partial x}, \end{aligned} \tag{4.2}$$

which tell that surface elevations and fluxes at each side of the steps match. Furthermore, due to the dispersion relation for shallow water waves,

$$\omega_i^2 = gk_1^2 h_1 = gk_2^2 h_2 = gk_3^2 h_1. \tag{4.3}$$

Note that the incident and transmitted waves have the same wavelength L_W. By combining Eq. 4.1 and Eq. 4.2, Mei et al. (2005) presented the solution for the amplitude T,

$$T = \frac{4s}{(1+s)^2 e^{-ik_2 W_r} - (1-s)^2 e^{ik_2 W_r}}, \tag{4.4}$$

where $k_2 = \frac{k_1 h_1}{s h_2} = \frac{2\pi}{L_W}\frac{h_1}{s h_2}$ and $s = \sqrt{h_1/h_2}$ for this problem. Note that the presented solution is valid as long as $h_1 k_1 << 1$ and $h_2 k_2 << 1$, according to shallow water

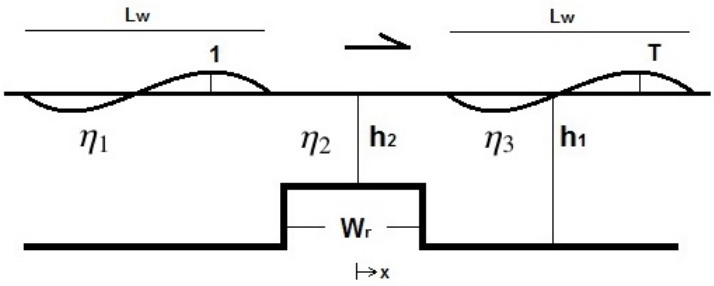


Figure 4.1: Wave transmission over a ridge with finite width W_r.

wave theory. It is important to remark also that the solution is based on linear theory. Thus, any incident linear wave can be considered, by decomposing it into sinusoidal waves of different wavenumber and amplitude.

Fig. 4.2 shows values of T, calculated with Eq. 4.4 and using different ratios $h_2/h_1 = 1/s^2$ and W_r/L_W. We observe that depth variations of 50% (i.e. $h_2/h_1 = 0.5$) reduce the transmitted wave amplitude in less than 5% (i.e. $T/1 > 0.95$). For smaller values of h_2/h_1 (i.e. a taller ridge), the reduction of the transmitted wave amplitude becomes relevant and depends on the ratio W_r/L_W.

The results of this analytical experiment can be used to draw two conclusions about the influence of the bathymetry uncertainty on the tsunami response uncertainty. First, bathymetry uncertainty (which can be related with the difference between h_1 and h_2) is relevant when it is comparable to the mean bathymetry (i.e. s is far from 1). Note that the bathymetry uncertainty can be represented by the uncertainty standard deviation. Thus, the ratio between the bathymetry uncertainty standard deviation and mean bathymetry (ratio known as coefficient of variation) needs to be close to unity so bathymetry can contribute to tsunami response uncertainties. It is important to remark that Mofjeld et al. (2001) studied the scattering of tsunami waves due to small scale features

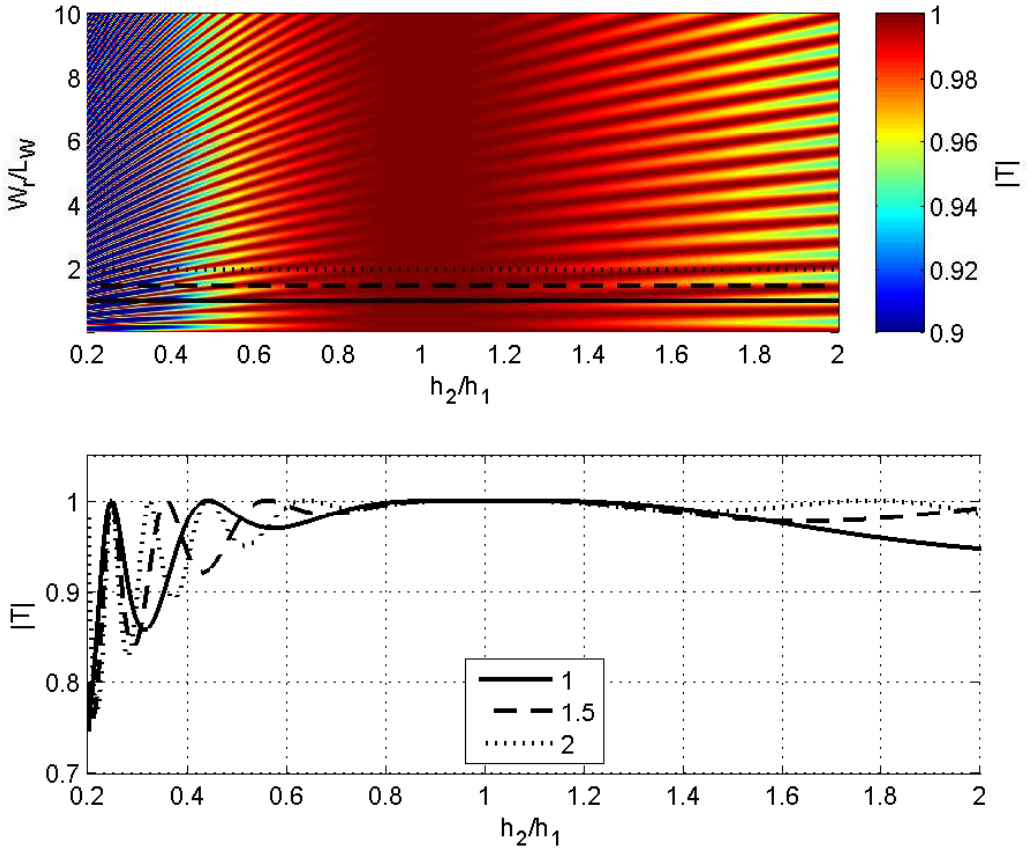


Figure 4.2: Amplitude of transmitted wave, T, after to propagate through a ridge with finite width W_r. The incident wave has unit amplitude. Top panel: Amplitude of transmitted wave for different values of h_2/h_1 and W_r/L_W. Bottom panel: Transects of the top panel for W_r/L_W=1, 1.5 and 2.

of the bathymetry. By using a numerical model, they simulated tsunami waves propagating over modified bathymetries of the Pacific Ocean, in which they removed real ridges and islands. They noted that the effect on tsunami waves is only relevant for high ridges, in which the depth is very shallow as compared with the surrounding ocean depth. Second, the tsunami uncertainty depends on the ratio between the horizontal length scales of the bathymetry uncertainty

(which can be related with W_r in the motivational experiment) and the tsunami wavelength. Note that the characteristic length of the bathymetry uncertainty is related with the spatial covariance. Based on these two conclusions, the assessment of tsunami uncertainties requires a consistent definition of the variance and spatial covariance of the bathymetry uncertainty. Furthermore, the stochastic methodology should generate consistent bathymetry samples for the uncertainty propagation.

4.2 Bathymetry uncertainty definition

In our methodology we first assume that uncertainties due to errors of surveyed data are negligible, as compared with interpolation errors. Thus, we only address bathymetry uncertainty due to interpolation errors. The bathymetry uncertainty is addressed by adopting a stochastic approach, in which bathymetry is modeled as a conditional random field. The construction of the conditional random field involves the construction of a prior unconditional random field, which is not constrained by bathymetry data. Then, a conditional random field is determined by combining the unconditional random field and surveyed bathymetry data in Eq. 2.36 and Eq. 2.37. The surveyed data, considered in this illustration case, is the surveyed data of GEBCO, which includes nautical charts and ship track soundings. The interpolated data of GEBCO was not considered in this study. In relevant coastal areas, we further include nautical charts of the Chilean Navy (SHOA, 2009).

The unconditional random field is assumed to be completely defined by the marginal distribution and covariance. The marginal distribution is unknown.

However, since uncertainties are unbiased and not bounded, we assume a Gaussian marginal distribution. The Gaussian distribution is defined by the mean and the variance, which are also not known. The mean of the bathymetry is assumed to be equivalent of a linear interpolation using the surveyed data. The left panels of Fig. 4.3 shows the linear interpolations within mesh 2 and mesh 3 using surveyed data, while the right panels show the surveyed data as red dots. The correlation function adopted in this illustration case is the Von Karman function, which was also adopted by Goff and Jordan (1988) to study spatial properties of the seafloor. The covariance matrix, associated with the Von Karman correlation function, is defined by the marginal distribution variance, the hurst number H and the correlation lengths a_x and a_y. Fig. 4.4 shows the power spectra of topo/bathymetry transects within the bathymetric grids employed in this thesis (North Chile and South China Sea). The transect locations were chosen to coincide with surveyed areas. The left panels are spectra of transects with East-West orientation, while the right panels are spectra of transects with South-North orientation. As stated by Bell (1975), all the spectra have a decay which is close to a -2 power law. The hurst number H, therefore, is set to 1, so the power spectrum at high wavenumbers has a -2 power law decay. Reasonable values for the variance and correlation lengths have not been determined in the literature so far. Thus, we estimate these values by means of sensitivity analysis and surveyed data.

The correlation lengths of the Von Karman function for the random bathymetry have been suggested to be smaller than 100 km (Bell, 1975). Their exact value, however, are not known. Kopp et al. (2016) published a high resolution bathymetry image within a portion of the application case domain in Northern Chile, which is shown in the top panel of Fig. 4.5. Bathymetry data,

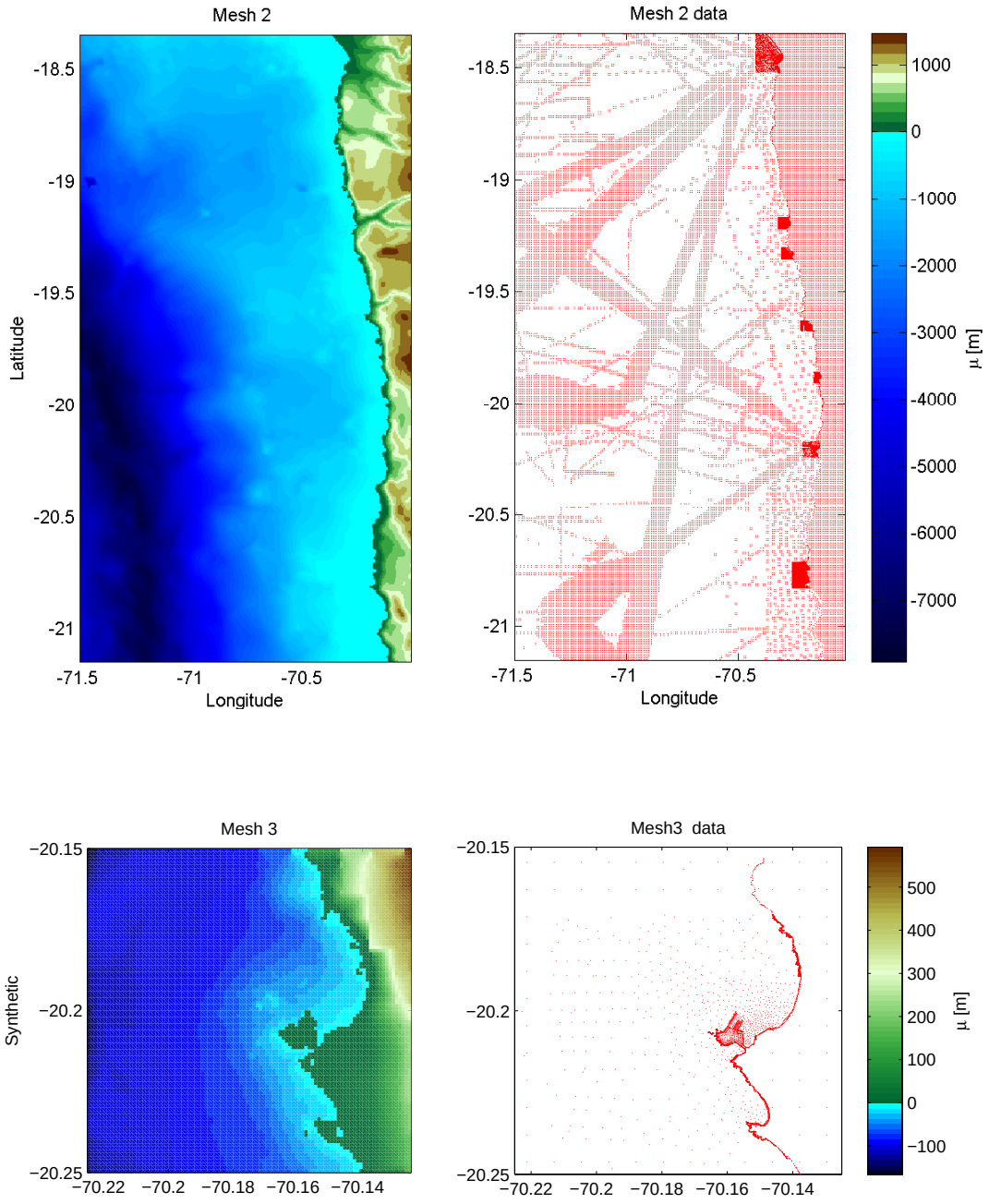


Figure 4.3: Left panels: Linear interpolation for mesh 2 and mesh 3 using bathymetry data. These interpolations are used as the mean bathymetry in unsurveyed regions. Right panels: Geographical location of surveyed data as red points. Data is obtained from ship tracks and nautical charts from Gebco, nautical charts from the Chilean Navy and topography satellite measurements.

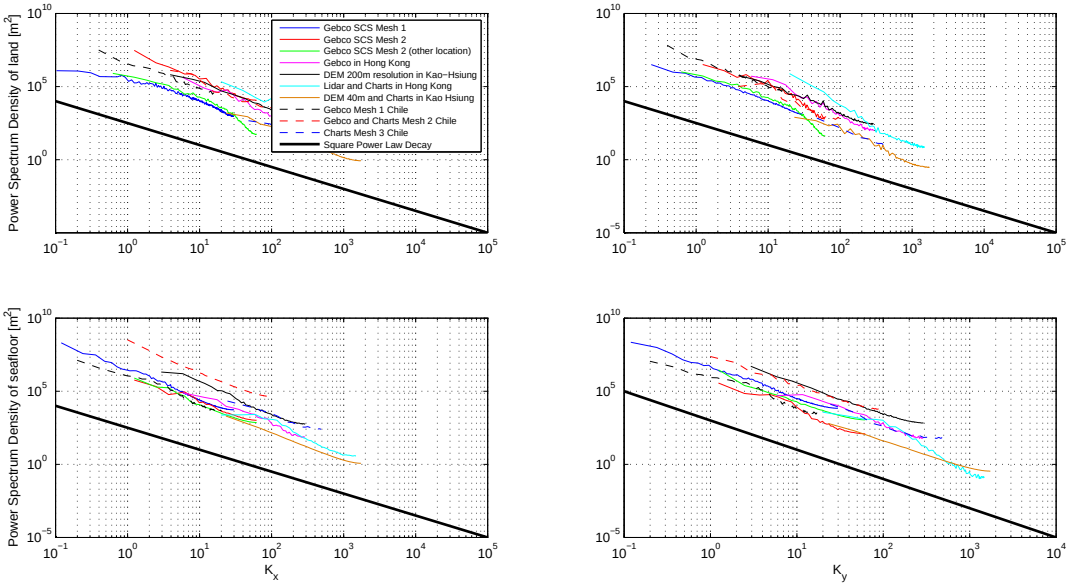


Figure 4.4: Spectral behavior of topo/bathymetry transects from the datasets used in this study. The unit of wavenumbers is $[^o]^{-1}$.

however, is not currently available. By means of a sensitivity analysis and visual comparison with the bathymetry image of Kopp et al. (2016), we specify reasonable values for the variance and the correlation lengths. The six bottom panels of Fig. 4.5 show the bathymetry samples of a portion of mesh 2 adopting two different correlation lengths, equal to $10km$ and $50km$. The variance σ_p^2 is also varied with values of $1.28\times10^5m^2$, $5\times10^5m^2$ and $2\times10^6m^2$. By visual comparison we have selected the random field with $a_x = a_y = 10km$ and $\sigma_p^2 = 5\times10^5m^2$ as the random field which better represent the bathymetry spatial variation in the surveyed data of Kopp et al. (2016). It is important to mention that high resolution bathymetry data is very scarce and often not available in tsunami hazard assessments. Hence other means have to be adopted to specify the correlation lengths and variance for other study cases. As an alternative, transects from ship track soundings can be used to determine the correlation lengths and the variance.

The conditional random field is constructed using the surveyed data con-

tained in GEBCO and Nautical charts. Further conditional points are also included at the boundaries of the meshes to allow a smooth transition between nested grids, without discontinuities. At the boundaries of the uncertain region of mesh 2 we specify a depth equal to the mean (i.e. no noise) to coincide with the depths of the assumed certain region of mesh 2, whose depths has been obtained from a linear interpolation. To avoid discontinuities at the boundaries of mesh 3, we also specified the depths to be equal to the depths in the uncertain region of mesh 2, which vary for each sample.

We remark that the methodology might be limited by computer capacity. The covariance matrix has a total of M_n^2 terms, where M_n is the number of grid nodes. For instance, mesh 2 has $M_n = 166352$ grid nodes, which means that its c_b would have 2.7×10^{10} terms. The required RAM memory to store that matrix would be $\sim$ 221 GB. The process to invert c_{p22} in Eq. 2.36 and to compute the eigenvalue decomposition of the K-L expansion of c_b are equally demanding. This thesis uses a computer with $\sim$ 60 GB of RAM memory, which allows to treat the bathymetry uncertainty of mesh 2 partially. We selected an uncertain region of mesh 2 with 291×291 grid nodes, which contains the region of tsunami generation and the assessed locations of Iquique (Iquique port and Iquique city in Fig. 3.3). Mesh 3, on the other hand, is entirely uncertain. It is important to mention that the tsunami response at offshore assessed locations (DART buoys in Fig. 3.3) or trailing tsunami waves in all the assessed locations might have previously propagated through shallow areas, which have not been specified as uncertain in this study. As a consequence, the tsunami uncertainty of those tsunami waves might not be well quantified. Future studies of the bathymetry uncertainty would require a bigger computer or numerical techniques to reduce the memory demand.

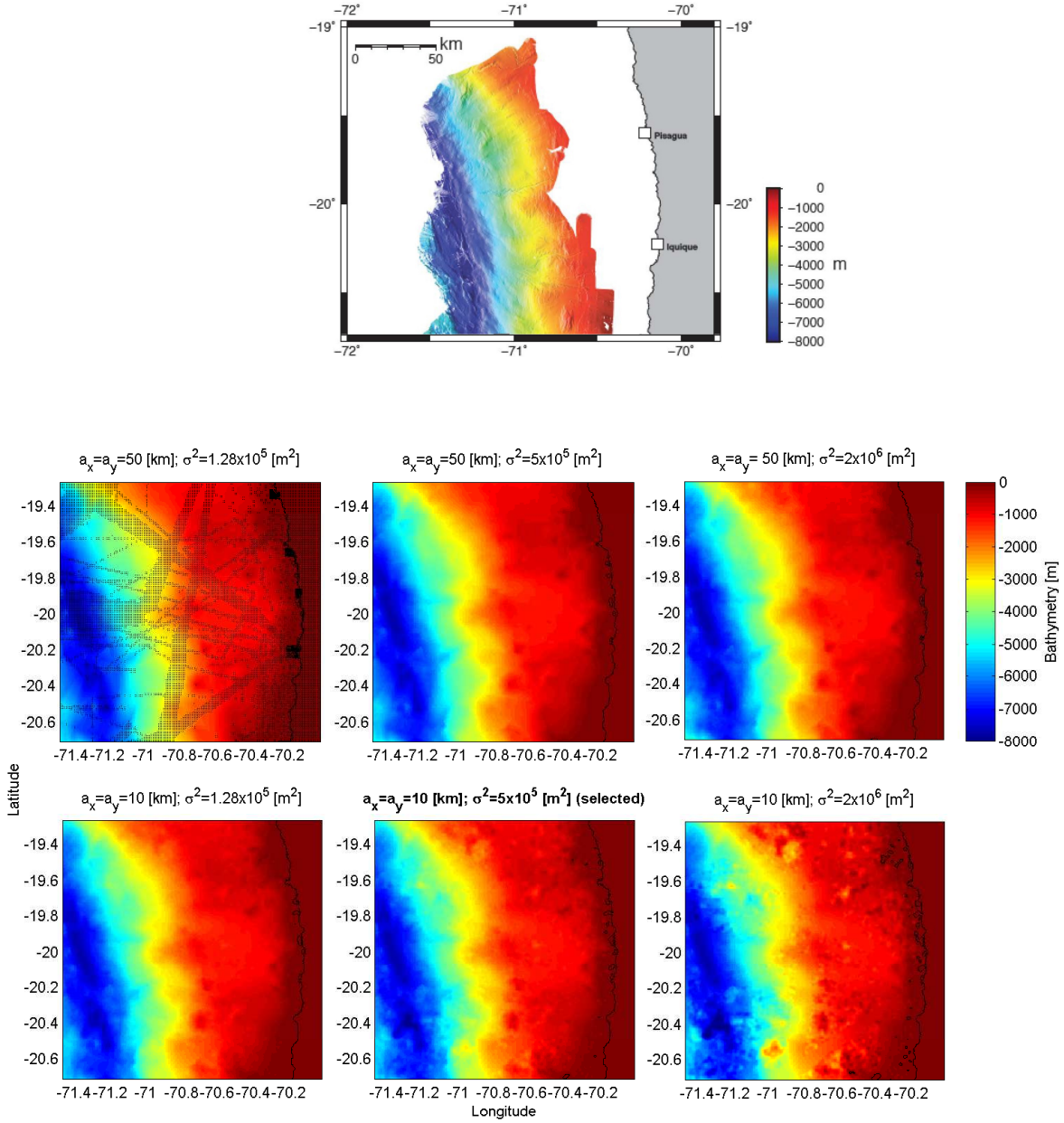


Figure 4.5: Top: High resolution bathymetry data surveyed by Geomar (Kopp et al., 2016). This data partially covers the study area of the application case. Bottom panels: Six Realizations of random fields with different correlation lengths a_x and a_y, and variance σ_p^2. The first realization shows surveyed data from GEBCO, nautical charts and satellite topography data as black dots. We adopt a random field with $a_x = a_y = 10km$ and $\sigma_p^2 = 5 \times 10^5 m^2$, which better replicates the spatial variability of the bathymetry data of the top panel.

4.3 Sample generation

The bathymetry conditional covariance c_b and conditional mean μ_b, given by Eq.2.36 and Eq.2.37, are used to generate bathymetry samples by means of a K-L expansion.

4.3.1 Accuracy of generated samples

In Section 3.3.1 we define a discretization error and truncation error associated with the generation of earthquake samples using a K-L expansion. In the uncertain bathymetry problem, the discretization of the bathymetry field (i.e. grid resolution and nested grids setup) is specified by the modeler according to physical arguments related with the characteristic length of tsunami waves. Thus, we only assess accuracies regarding to the truncation error, which was defined in section 2.4.2.

The truncation error for the uncertain bathymetry problem is defined as the mean loss of variance among the grid nodes of the bathymetry, as a result of the truncation. The truncation error for the $i-th$ grid node $\epsilon_{T,b(i)}$ is given as,

$$\epsilon_{T,b(i)} = \frac{\Psi_{N(i)}\Lambda_N \left(\Psi_{N(i)}\right)^T - \sigma^2_{b(i)}}{\sigma^2_{b(i)}} \tag{4.5}$$

where $\Psi_{N(i)}$ corresponds to the $i-th$ row of the eigenvector matrix of c_b, with N eigenvectors as columns. Λ_N is a diagonal matrix with N eigenvalues and $\sigma^2_{b(i)}$ is the bathymetry variance of the $i-th$ grid node, which is equal to the $i-th$ diagonal term of c_b. Fig. 4.6 shows the average loss of variance among the grid nodes of meshes 2 and 3 in terms of different truncations. In the illustration case we consider 5500 K-L terms for mesh 2, which are associated with an av-

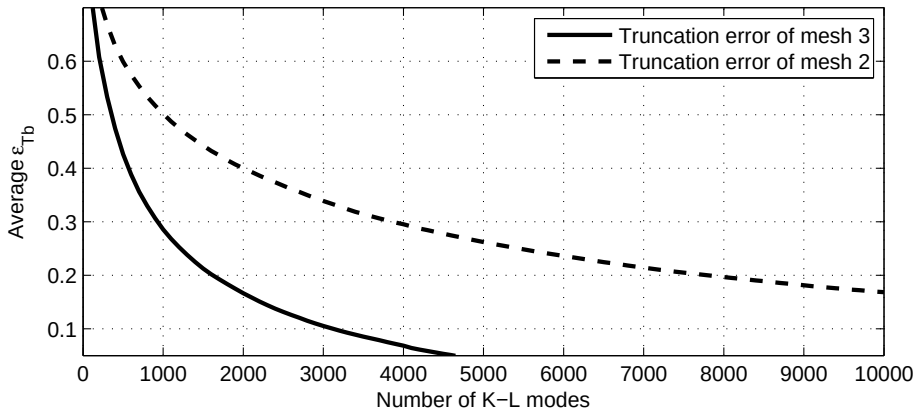


Figure 4.6: Truncation error ϵ_{Tb} in terms of number of K-L modes.

erage truncation error smaller than 25%. For mesh 3 we use 1500 K-L terms, which are associated with an average truncation error smaller than 23%. It is important to mention that the adoption of smaller truncation errors (i.e. more K-L terms) have two important impacts on the computational cost. First, the eigenvalue decomposition of c_b would require a longer time to compute more eigen-modes (K-L modes). Second, the optimization of SROM of below would have more dimensions and, therefore, the optimization would require a longer time to find the optimal set of m samples. In this illustration case, the eigenvalue decomposition required 12 hours in a conventional computer. The solution of the optimization problem for the SROM required 48 hours.

4.3.2 Analysis of generated samples

The top panels of Fig. 4.7 shows 3 samples of the uncertain bathymetry within mesh 2, which are enclosed by segmented black lines. Note that the transition between the certain and uncertain bathymetry of mesh 2 is smooth, without dis-

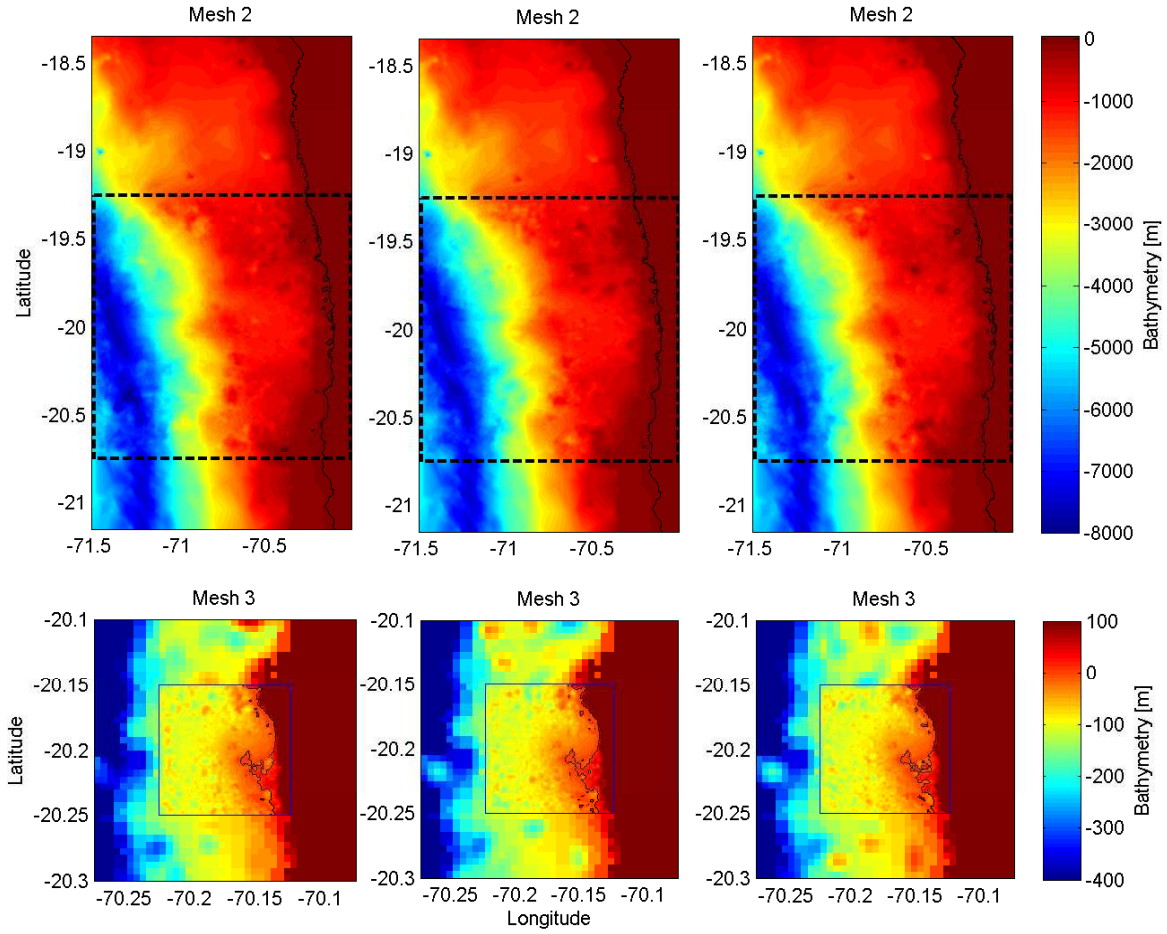


Figure 4.7: Three realizations of uncertain bathymetry random fields.

continuities. The bottom panels of Fig.4.7 shows three samples of mesh 3, which are enclosed by blue lines. We have also plotted the bathymetry of mesh 2 to show the smooth transition between meshes 2 and 3. It is important to mention that mesh 3 seems to have a spatial variability with smaller length scales. This is a consequence of the finer grid resolution of mesh 3.

4.4 Uncertainty Propagation

We adopt the SROM to propagate uncertainties from bathymetry to tsunami responses using m=200 bathymetry samples. After we select the bathymetry samples, COMCOT tsunami model is employed to assess the uncertainties in tsunami responses.

4.4.1 SROM of bathymetry

The m=200 bathymetry samples $\tilde{Z}$ are selected by solving an optimization problem, in which the statistics of the 5500 K-L coefficients of mesh 2 samples and the 1500 K-L coefficients of mesh 3 samples are compared with the target probability properties. The optimization problem considers the discrepancies of distributions, second moments properties and covariance,

$$\begin{aligned} e_1(\tilde{Z}, \vec{p}) &= \sum_i \int_{I_i} (\tilde{F}(x_i) - \Phi(x_i))^2 dx_i \\ e_2(\tilde{Z}, \vec{p}) &= \sum_i \sum_{q=1}^{2} \left(\tilde{\mu}_i^q - \mu_i^q\right)^2 \\ e_3(\tilde{Z}, \vec{p}) &= \sum_{i,j;j>i} \left(\tilde{c}_{i,j} - c_{i,j}\right)^2 \end{aligned} \tag{4.6}$$

where I_i is the range of possible values of the $i-th$ K-L coefficient, $\Phi(x_i)$ and $\tilde{F}(x_i)$ are the cdf's of the $i-th$ K-L coefficient, associated with Z (which is Gaussian) and $\tilde{Z}$, respectively. μ_i^q and $\tilde{\mu}_i^q$ are the mean (q=1) and variance (q=2) of the $i-th$ K-L coefficient, associated with Z (equal to zero mean and unit variance) and $\tilde{Z}$, respectively. $c_{i,j}$ and $\tilde{c}_{i,j}$ are the covariance of the i-th and j-th K-L coefficients, associated with Z (equal to an identity matrix) and $\tilde{Z}$, respectively. The values for the weights α_1, α_2 and α_3 in Eq. 2.43 can be obtained by trial and error so

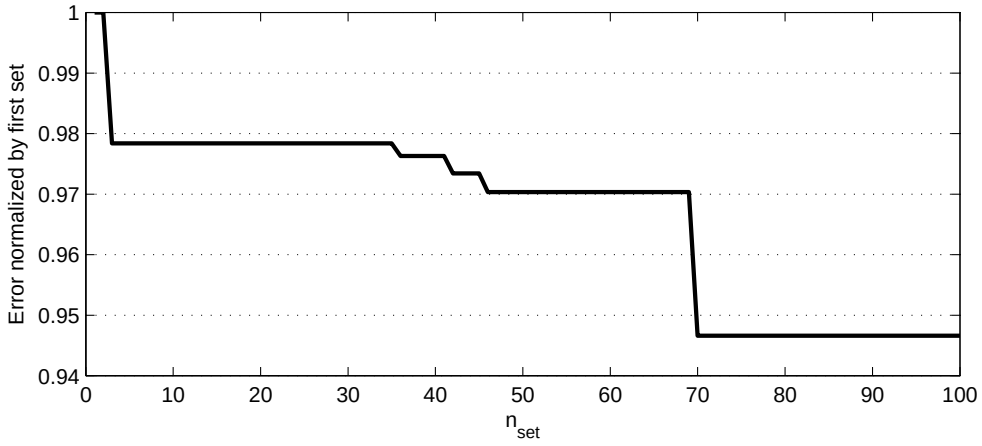


Figure 4.8: Discrepancies between the $m = 200$ samples and the target probability properties in terms of n_{set} of the sub-optimal procedure. Discrepancies are normalized by the discrepancy of the first tested set.

that the contributions of e_1, e_2 and e_3 in e_t have the same order of magnitude. For this specific illustration case we use $\alpha_1 = 1.0$, $\alpha_2 = 28.5$ and $\alpha_3 = 32$. As the case of Chapter 3, we again use the MATLAB function **fmincon** with an interior point algorithm to solve the optimization problem. We also adopt the sub-optimal procedure of Grigoriu (2009). As it was described in Section 2.4.3, this procedure randomly generates n_{set} sets of m=200 samples. An optimization problem is then solved for each set, which only considers the probabilities as variables. Fig. 4.8 shows the reduction of error in terms of n_{set}.

4.4.2 SROM of tsunami model output

We use COMCOT tsunami model and the same model setup described in Chapter 3 to run m=200 tsunami simulations with different bathymetry samples for mesh 2 and mesh 3. To assess the effect of bathymetry only, we use the same

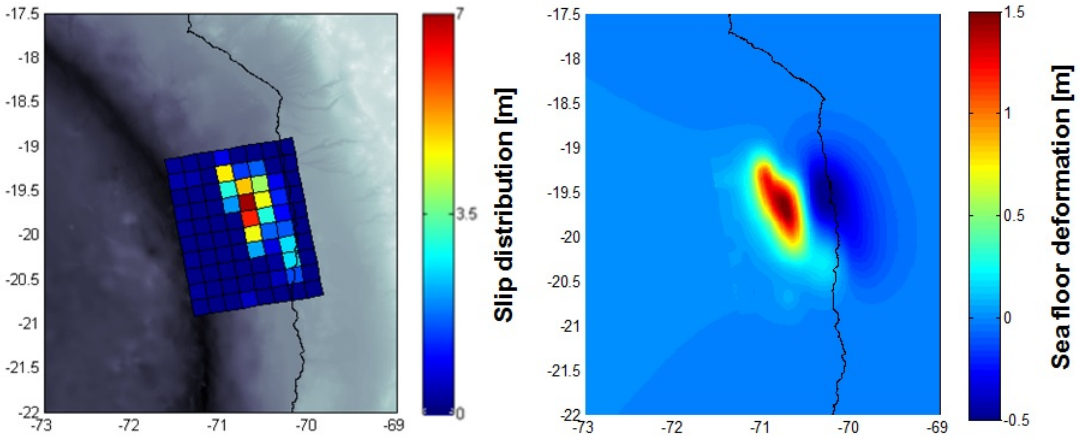


Figure 4.9: Earthquake slip distribution (left panel) and tsunami initial condition (right) proposed by An et al. (2014) for the 2014 Chilean Earthquake. The slip distribution was obtained by solving an inversion problem using tsunami records.

earthquake scenario in all the m=200 tsunami simulations. We consider the fault parameters of the 2014 Chilean earthquake, which were estimated by An et al. (2014). This reference used DART buoy records to estimate the slip distribution of the earthquake. Fig. 4.9 show the estimated slip distribution and tsunami initial condition associated with the earthquake.

Fig. 4.10 shows time histories of the simulated tsunami amplitudes at different locations around the seismogenic region: two in the front of the port (Lon-70.148^o Lat-20.204^o) and city of Iquique (Lon-70.170^o Lat-20.220^o) and two at offshore buoys locations (DART32401 at Lon-73.429^o Lat-20.473^o and DART32402 at Lon-73.983^o Lat-26.743^o). We also assess the tsunami at three additional locations where uncertainties have been observed to be significant. These correspond to the locations $A1$ (Lon-70.150^o Lat-21.425^o), $A2$ (Lon-70.440^o Lat-22.960^o) and $A3$ (Lon-71.250^o Lat-17.850^o), which are shown in Fig. 4.12. The time histories of the earthquake samples are depicted as black

curves. As the case of Chapter 3, we have also computed the 50%, 15% and 5% exceedance probabilities of the time histories.

4.4.3 Analysis of uncertainty propagation

By using the maximum tsunami amplitudes of the samples of Fig. 4.10, we build probability exceedance curves, which are shown as black curves in Fig. 4.11. We also show the coefficient of variation (CoV) in the headers of each panel, which is a measure of the strength of the tsunami uncertainty. As it is observed in Fig. 4.10, the two assessed locations in Iquique have small uncertainties in their leading waves. This is a consequence of small uncertainties in shallow waters, where bathymetry uncertainties influence tsunami responses. The maximum tsunami amplitude in Iquique, however, is significantly uncertain, as shown in Fig. 4.11. This is explained by the fact that trailing waves, which contain the maximum tsunami amplitudes in Iquique, are significantly uncertain due to reflected waves coming from shallow areas, in which bathymetry is more uncertain. The time histories at $A1$, $A2$ and $A3$ in Fig. 4.10 show that maximum tsunami amplitudes are also associated with trailing waves, which have previously propagated over uncertain shallow waters. The observations at these three additional locations suggest that edge waves and resonant effects can be greatly affected by bathymetry uncertainties. The uncertainties of the maximum tsunami amplitude at the DART bouys are the smallest among the assessed locations. The small uncertainties are the result of leading waves containing the maximum amplitudes. Since leading waves do not propagate over uncertain shallow waters before they arrive to the DART buoys locations, their uncertainties are very small.

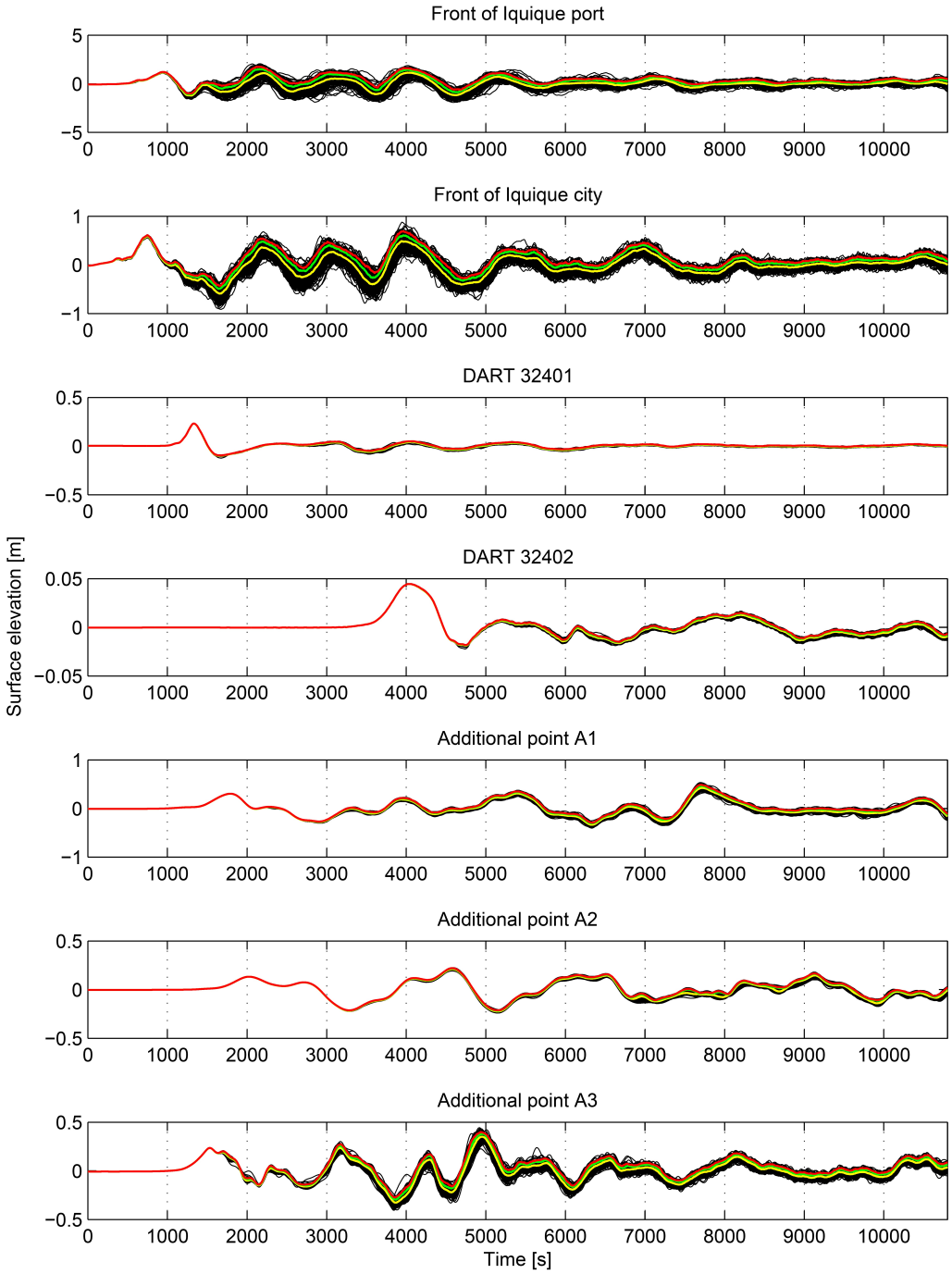


Figure 4.10: Tsunami wave elevations at seven locations close to the rupture area. Black time histories corresponds to m=200 tsunami samples with consideration of uncertain bathymetry. The red, green and yellow curves indicate the 5%, 15% and 50% exceedance probabilities, respectively.

In the motivation experiment of section 4.1 we conclude that the bathymetry uncertainty standard deviation and covariance play an important role in tsunami uncertainty assessments. To assess the pertinence of these probability properties in the 2014 illustration case, we further perform a sensitivity analysis, in which we have reduced the unconditional bathymetry variance to $\sigma_p^2 = 1.28 \times 10^5\ m^2$ (red curves in Fig. 4.11) and we have increased the correlation lengths to $a_x = a_y = 50km$ (blue curve in Fig. 4.11). The reduction of variance is related with a reduction of uncertainty, while the increase of correlation lengths is related with a smoother spatial variability of the bathymetry samples. Note that the curves are significantly different with respect to the black curves (with $\sigma_p^2 = 5 \times 10^5\ m^2$ and $a_x = a_y = 10km$). This analysis also demonstrates that tsunami assessment uncertainties are very sensitive to the variance and correlation lengths of the bathymetry uncertainty.

To better understand the effect of bathymetry uncertainties on maximum tsunami amplitudes, Fig. 4.12 shows the coefficient of variation (CoV, equal to standard deviation divided by the mean) of the maximum tsunami amplitude for each grid node of mesh 1. This plot was obtained by calculating the ensemble average and standard deviation of the maximum elevation in mesh 1 from the m=200 tsunami simulations. Fig. 4.12 shows that the largest CoV's are located in shallow areas, which are close to the uncertain bathymetry area (depicted as a red square). The smallest uncertainties are located inside, at the south and at the southwest of the tsunami generation zone. These areas have been observed to have leading waves containing the maximum amplitude and which have not previously propagated over uncertain shallow areas. An interesting conclusion from Fig. 4.12 is that assessed locations within some offshore regions are not significantly affected by bathymetry uncertainties. As a consequence, inversion

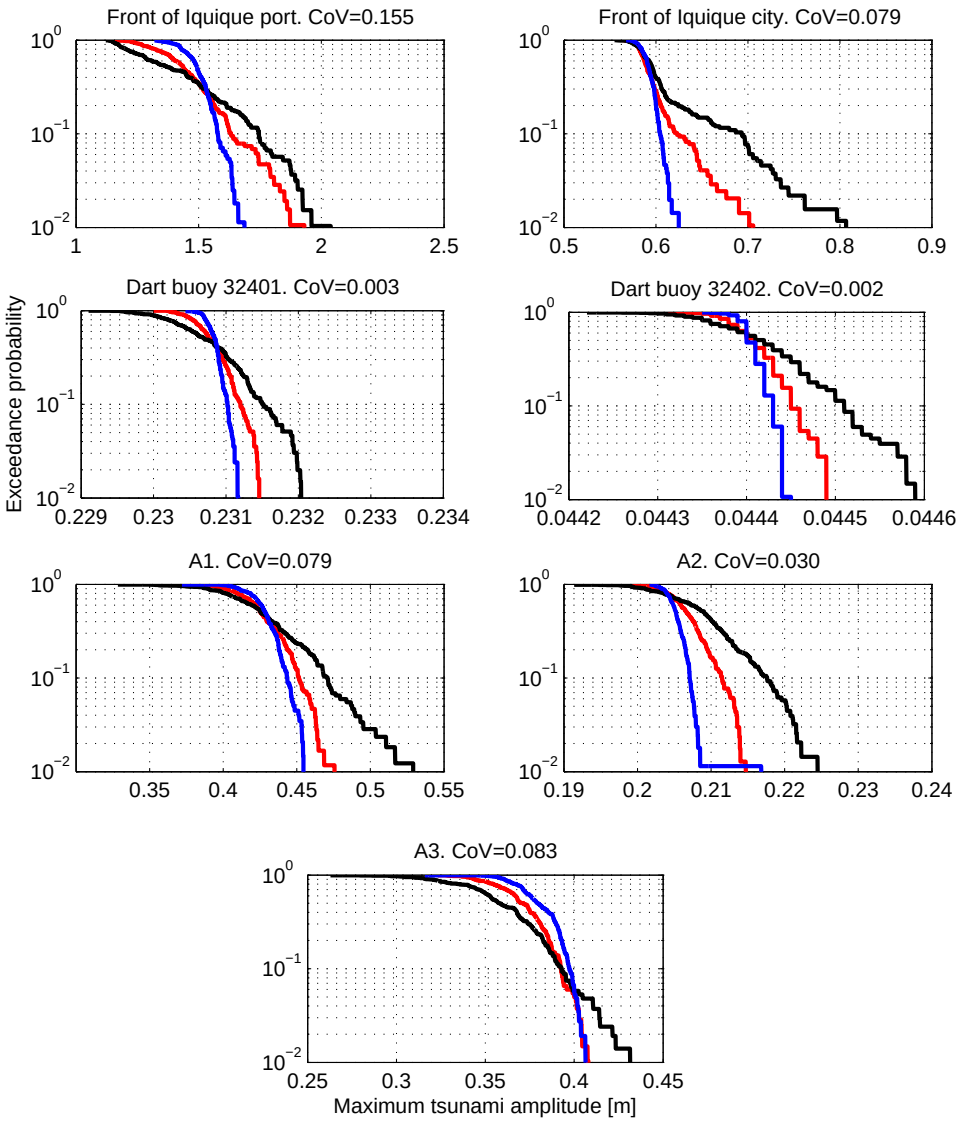


Figure 4.11: Exceedance probability curves of the maximum tsunami amplitude relative to the still water level at the seven assessed locations, estimated with an SROM with m=200 bathymetry samples. Black curves: Exceedance probability curves corresponding to the uncertain bathymetry model adopted in this thesis, with $\sigma_p^2 = 5 \times 10^5 m^2$ and $a_x = a_y = 10km$. Red curves: Exceedance probability curves using $\sigma_p^2 = 1.28 \times 10^5 m^2$ and $a_x = a_y = 10km$. Blue curves: Exceedance probability curves using $\sigma_p^2 = 5 \times 10^5 m^2$ and $a_x = a_y = 50km$.

methods based on some DART buoy records, which are used to characterize earthquake fault parameters, are not severely affected by bathymetry uncertainties. Inversion methods using tsunami records at the coast, conversely, might be significantly affected by bathymetry uncertainties.

As in the case of Chapter 3, we compare our uncertainty quantification results with the records of the tsunami generated by the 2014 Chilean earthquake. Similar comparisons were also conducted in An et al. (2014), who used an assumed deterministic bathymetry based on GEBCO (surveyed and interpolated data) and nautical charts from the Chilean Navy. As it is expected, Fig. 4.13 shows small discrepancies between simulations and tsunami records (magenta line) at the DART buoys. This because the earthquake slip distribution was estimated using an inversion method based on these DART buoys. The comparison in Iquique port, however, shows large discrepancies in amplitudes and phases. These discrepancies are similar to those obtained in An et al. (2014). This comparison suggests that bathymetry uncertainties play a secondary role in the uncertainties of tsunami assessments. We also observe that bathymetry uncertainties do not explain the discrepancies of amplitudes and phases of trailing waves, which are observed in the results of Chapter 3. Thus, other sources of uncertainty have to be analyzed.

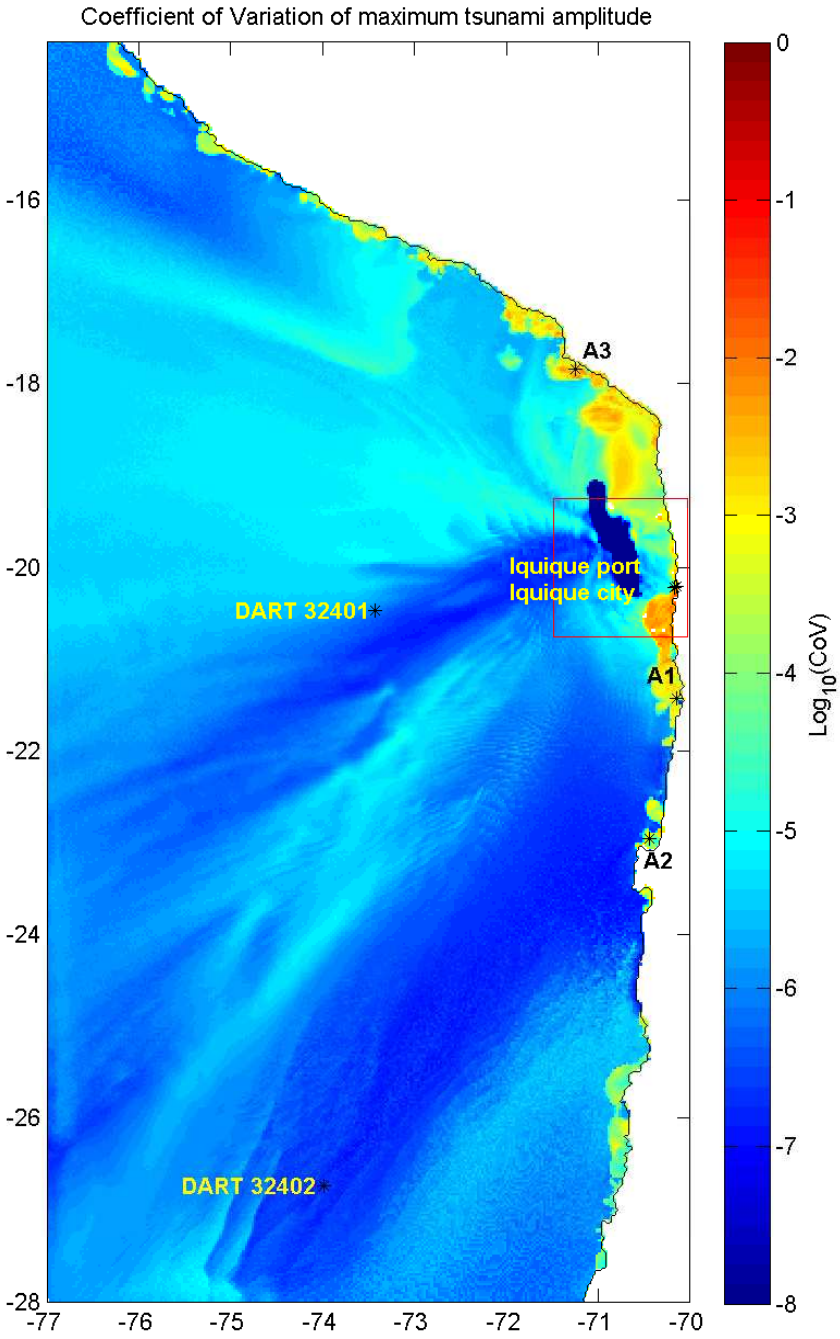


Figure 4.12: Coefficient of variation of the maximum tsunami amplitude over the tsunami propagation model domain. Red square indicates the area of uncertain bathymetry. Black stars indicate the assessed locations.

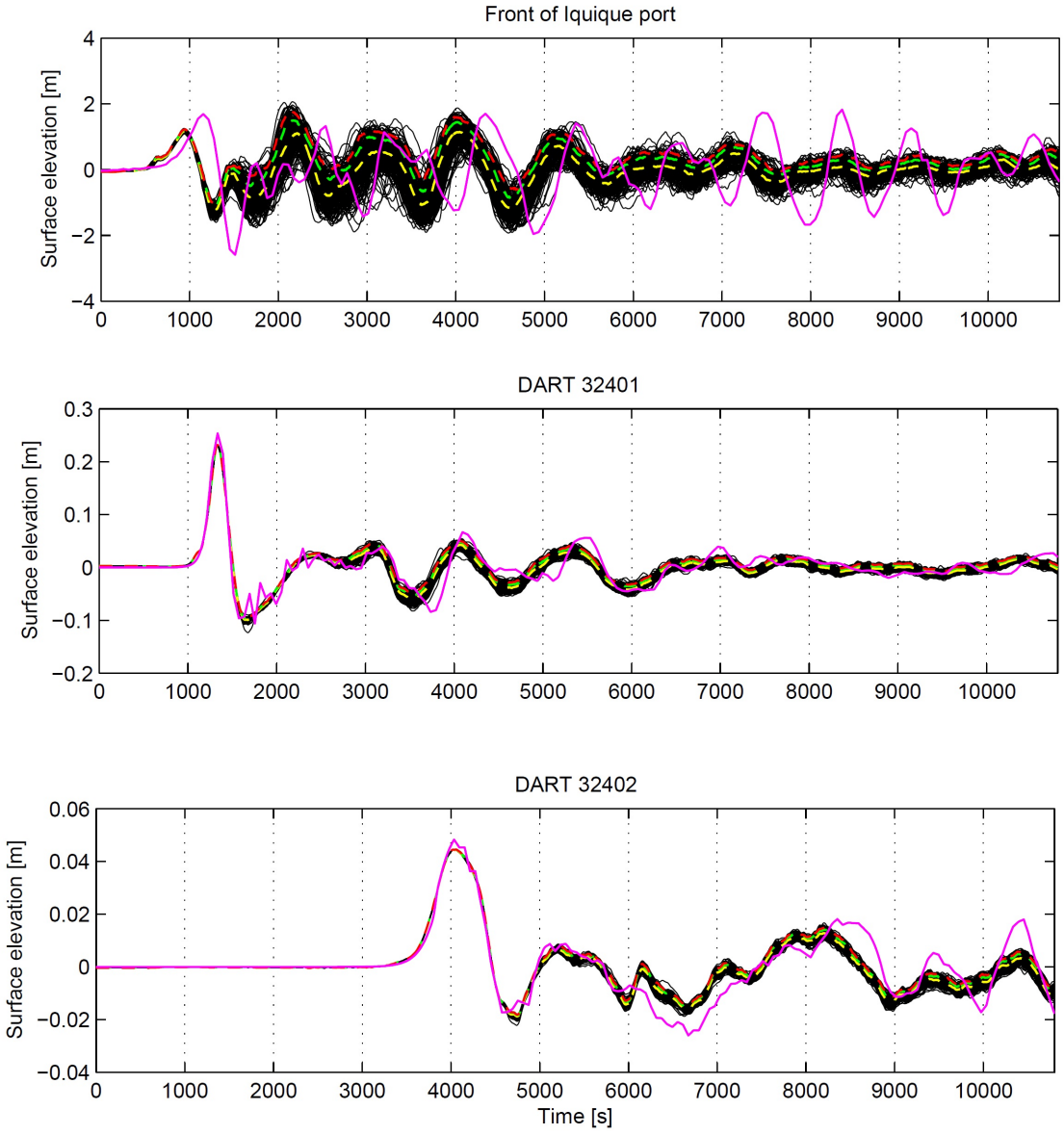


Figure 4.13: Comparison of the 200 SROM samples as black curves and measurements of the tsunami in 2014 as magenta curves. The red, green and yellow dashed curves correspond to the exceedance curves of Fig. 4.10.

CHAPTER 5

PTHA IN SOUTH CHINA SEA

In this chapter we conduct a probabilistic tsunami hazard assessment (PTHA) in the South China Sea (SCS), with consideration of uncertain earthquakes triggered within the Manila Subduction Zone. The PTHA is focused in Hong Kong, China, and Kao Hsiung, Taiwan (Figure 5.1), where two of the most important ports in East Asia are located (Yap et al., 2006). Since the bathymetry data collected for these assessed locations is sufficiently accurate, we do not consider bathymetry uncertainties.

Hong Kong is a politically autonomous territory located in the Pearl River delta, China. Its population was estimated in 7,219,700 inhabitants in 2016 (Roux, 2016), which ranks this city as one of the most populated areas in the region. Hong Kong is the fourth most important financial city (Brunn et al., 2016) and the fourth most important container port (Weeks et al., 2017) worldwide. Kao Hsiung, on the other hand, is located in the south west coast of Taiwan and is the second most important city of the country, with a population of 2,778,918 inhabitants in 2016. Kao Hsiung port is located next to the city. This latter is the base of the Taiwanese Navy and the most important port of the country (Haynes et al., 1997). Hong Kong and Kao Hsiung ports are exposed to different natural marine hazards. Given the importance of both ports to the local and global economy, the quantification of their hazards and vulnerabilities is relevant.

According to Terry et al. (2017), the two most relevant vulnerabilities along the coasts of the SCS are the inundation due to tropical cyclones and tsunamis. This latter have been addressed extensively in the last decade (e.g. Li et al. (2016); Thio et al. (2007); Liu et al. (2007); Okal et al. (2011); Liu et al. (2009)).

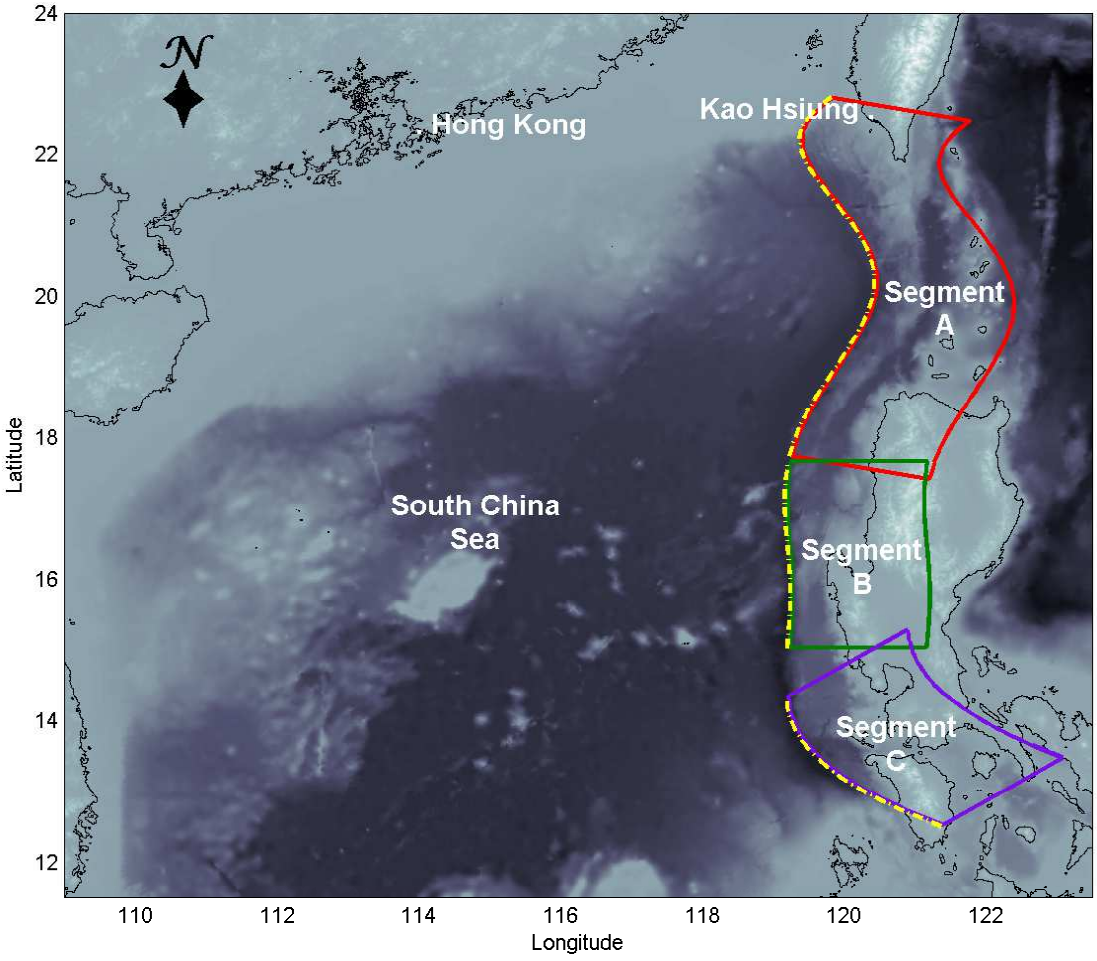


Figure 5.1: Map of the South China Sea and the coastal areas where the PTHA is assessed. The map also shows the hypothetical segmentation of the Manila Subduction Zone according to Li et al. (2016) and adopted in the present assessment. The yellow segmented line indicates the trench of the Manila Subduction Zone.

According to Terry et al. (2017), however, the tsunami hazard has not been well understood because of the lack of records of past tsunamis, the scarcity of paleotsunami studies and the lack of information to model the most relevant tsunami

sources.

Since the SCS is sheltered by islands and submarine mountain chains, the most hazardous tsunami sources along the coast are located inside this sea (Okal et al., 2011). Terry et al. (2017) identified three different types of tsunami sources within SCS: earthquakes, submarine landslides (collapse of clastic accumulation on continental shelf margins or carbonate build up) and volcanic related sources (eruptions or flank collapses of islands and seamonts). Earthquakes have been identified as the most hazardous source due to the presence of active subduction zones. The Manila Subduction Zone, in particular, is the subduction zone with the potential to host great earthquakes ($M_W > 8.5$) and to generate the largest tsunamis in SCS (Terry et al., 2017; Okal et al., 2011; Liu et al., 2009; Megawati et al., 2009). Thus, the PTHA presented in this chapter considers tsunamis generated by earthquakes in the Manila Subduction Zone. It is important to mention, though, that Terry et al. (2017) suggested that other type of sources, generating smaller tsunamis, may be associated with significant tsunami waves close to the source. For instance, they mentioned that Hong Kong may be affected by near field landslide tsunamis and Kao Hsiung by near field landslide and volcanic-related tsunamis. These type of sources and their corresponding tsunami hazards have not been investigated in the literature so far and should be included in the future to assess the tsunami hazard of Kao Hsiung and Hong Kong completely.

The PTHA aims to assess the maximum tsunami amplitude and the inundation in Hong Kong and Kao Hsiung ports in terms of their probabilities in a period of time. The tsunami amplitudes are assessed at different locations, which are shown by numbers in Fig. 5.2 and Table 5.1 (stations 1 to 6). We

Database	Lon. [^{o}]	Lat. [^{o}]	Vertical accuracy
Assessed locations of this study			
Station 1	120.2800	22.5700	Offshore side in Kao Hsiung Port.
Station 2	120.2750	22.6130	North side, inside Kao Hsiung Port.
Station 3	120.3290	22.5460	South side, inside Kao Hsiung Port.
Station 4	114.1200	22.1000	South Lamma Island, South Hong Kong.
Station 5	114.2578	22.2701	East Harbour entrance to Hong Kong.
Station 6	114.2030	22.2998	Front of Kai Tak Terminal of Hong Kong.
Additional locations for lineal model in coarse grid (Section 5.3)			
Station 7	114.3675	22.3102	North of Hong Kong.
Station 8	114.0350	22.1869	West of Lamma Island, South of Hong Kong.
Station 9	113.8250	22.1096	South of Guishanzhen Island, South Hong Kong.
Station 10	120.3510	22.5010	Offshore, north Kao Hsiung.
Station 11	120.2200	22.6900	Offshore, south Kao Hsiung.
Station 12	120.2900	22.5420	Offshore, front Kao Hsiung.

Table 5.1: The six assessed locations of this study (station 1 to 6) and six additional locations used to compare this study with existing PTHA assessments. Stations 7,8 and 9 corresponds to the Hong Kong assessment locations of Li et al. (2016).

assess two unsheltered locations (station 1 and station 4), where we expect to see larger tsunami amplitudes, and four sheltered locations (stations 2, 3, 5 and 6), where most of the port activities take place. In Table 5.1 we also present 6 additional stations (station 7 to 12), which are used for a comparison with existing PTHA studies (section 5.3). The inundation, on the other hand, is assessed spatially in Hong-Kong and Kao-Hsiung. The assessed regions correspond to grids 5a and 5b in Fig. 5.6.

5.1 Definition of synthetic earthquakes in Manila Subduction Zone

The tsunamis studied in this PTHA are originated in the Manila Subduction Zone, in which the Sundaland (Eurasia) plate and the Philipines Sea plate con-

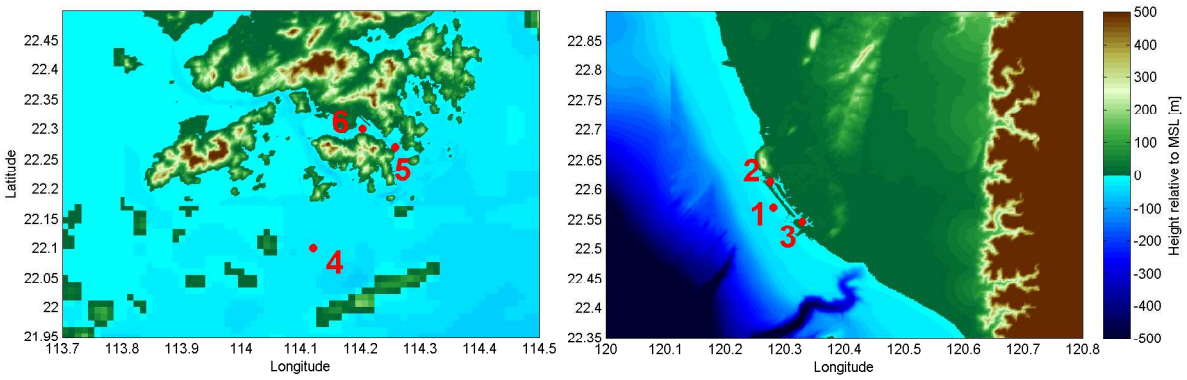


Figure 5.2: Detail of assessed locations in Hong Kong (left panel) and Kao Hsiung (right panel). Heights are relative to still water level (SWL).

verge. The subduction rate has been estimated in ~ 8 [cm/yr] (Megawati et al., 2009). Despite the large convergence rate (comparable to Chilean subduction zone rates), earthquakes of magnitude $M_W > 8$ have not reported in the literature. Some studies (Galgana et al., 2007; Lin, 2015) suggest that, despite the high subduction rate, the plate interface of the Manila Subduction Zone is partially coupled. Hence, these studies suggest that the potential of this subduction zone to generate large earthquakes is small. However, these studies use incomplete information and, therefore, the likelihood of large earthquake is still unknown (Megawati et al., 2009). The PTHA of this chapter considers that large earthquakes can be generated in the Manila Subduction Zone, as suggested by Hsu et al. (2012).

5.1.1 Geometry of the seismogenic regions.

According to the seismicity study of Hsu et al. (2012) and Li et al. (2016), the Manila Subduction Zone is divided in three segments, which are limited by the ends of the subduction zone and the Scarborough seamount chain. These segments are shown in Fig. 5.1. Based on spatial constraint considerations and the scaling relation of the rupture area in terms of earthquake magnitude of Blaser et al. (2010) (see Fig. 2.2), we estimate the largest earthquake magnitudes as 9.0, 8.5 and 8.5 for segment A, B and C, respectively.

The up-dip extent of the sismogenic region of these segments seems to be not well understood. However, it has been observed that the accretionary wedge (i.e. sediments on the trench) has a thickness of about 4-5 km (Hsu et al., 2012; Ludwig, 1970; Hayes and Lewis, 1984). In this study we assume that earthquakes can propagate up to a depth of 1 [km] in the accretionary wedge, below the trench. Hence, the up-dip extent of the seismogenic region is specified as 1 km. According to Hsu et al. (2012), the northern segments A and B have a dip angle of $\sim 15^o$ from the seafloor to 30 km depth. Between the depths 30 km and 89 km, a dip angle of 30^o is estimated, covering a total width of 230 km (Hsu et al., 2012) for the seismogenic regions. The southern segment C has the same dip angles of segments A and B for depths shallower than 50 km. However, for deeper depths segment C steepens to a dip angle of 50^o. The bottom depth of the seismogenic region of segment C is 110 km, according to a seismogenic region width of 230 km (Hsu et al., 2012). For sake of simplicity, we specified constant strike angles for segments A, B and C, equals to 10^o, 0^o and 330^o, respectively. The rake angles have a conservative value of 90^o.

5.1.2 Recurrence of earthquakes.

Different earthquake magnitudes are considered in a PTHA. By assuming that earthquakes in Manila Subduction Zone are well described by the truncated Gutenberg Richter Law (G-R), the mean recurrence of earthquakes with different magnitudes can be determined. The G-R law states that the number of earthquakes greater or equal to a certain magnitude follows the following expression,

$$N(M_W) = 10^{a-bM_w}, \tag{5.1}$$

where $N(M_W)$ is the number of earthquakes equal or greater than M_W per year and a and b are site specific parameters. Note that the reciprocal of $N(M_W)$ is equal to the return period in years of earthquakes equal or greater than M_W. In general, two approaches can be adopted to determine the a and b values of a seismogenic region. The first approach uses statistics of past earthquakes to calibrate Eq. 5.1. Fig. 5.3 shows the recurrences of past earthquakes withing each segment of Fig. 5.1 as black diamonds. The earthquake data was obtained from the catalog of the National Earthquake Information Center of the United States Geological Survey (NEIC) between the years 1973-2017. The blue lines, on the other hand, correspond to the curve fitting of Li et al. (2016) using the information of NEIC catalog (i.e. data similar to the black diamonds). Note that the b values correspond to the line slopes in the graphs, while the a values correspond to the intersection of the y axis at magnitude zero. The earthquake recurrence model using this first approach is defined as the seismic based approach herein. A second approach uses inter-seismic information to infer the earthquake recurrence. By using measurements of the rate of ground deformation close to the Manila Subduction Zone, Hsu et al. (2012) estimated the inter-seismic rate of accumulation of seismic moment, $\dot{M}_0$, in each of the three seismogenic regions.

This rate of accumulation is interpreted as the accumulation of energy which is released during earthquakes. According to Molnar (1979) and Ader et al. (2012), $\dot{M}_0$ is related with the a value of the G-R curve by,

$$a = Log_{10}\left(\frac{(1-2/3b)\dot{M}_0}{M_{0,Max}^{1-2/3b}}\right) - 6.07, \tag{5.2}$$

where b is assumed to be one and $M_{0,Max}$ is the seismic moment of the greatest earthquake within the assessed seismogenic region. Li et al. (2016) estimated a values for the three seismogenic regions, which are used to build the red G-R curves in Fig. 5.3. The earthquake recurrence model using this approach is defined as the geodetic based approach herein. In general, lower earthquake recurrences are observed with the seismic based approach for earthquakes with $M_W > 7$. According to Li et al. (2016), both approaches can be interpreted as extreme bounds of the true earthquake recurrence of the three segments. It is important to remark, though, that both approaches are subject to errors. For instance, the seismic based approach only uses statistics of 42 yr, while the amount of ground deformation data of the geodetic based approach is insufficient to accurately estimate the accumulation of seismic moment (Hsu et al., 2012). We remark here that the errors in the estimation of earthquake recurrences constitute an additional source of uncertainty in the PTHA. This additional source of uncertainty, though, is not further analyzed in this thesis. Hence, we adopt the same a and b values provided by Li et al. (2016) which are also presented in the headers of each segment in Table 5.2.

Since it is not possible to obtain the recurrence of a specific earthquake magnitude, we adopt discretization intervals $M_{Wj} \pm \frac{\Delta M_{Wj}}{2}$, as defined in section 2.3. The yearly recurrence of each magnitude interval ($\lambda^{EQ}_{M_{Wj},x_i}$ in Eq. 2.25), can be

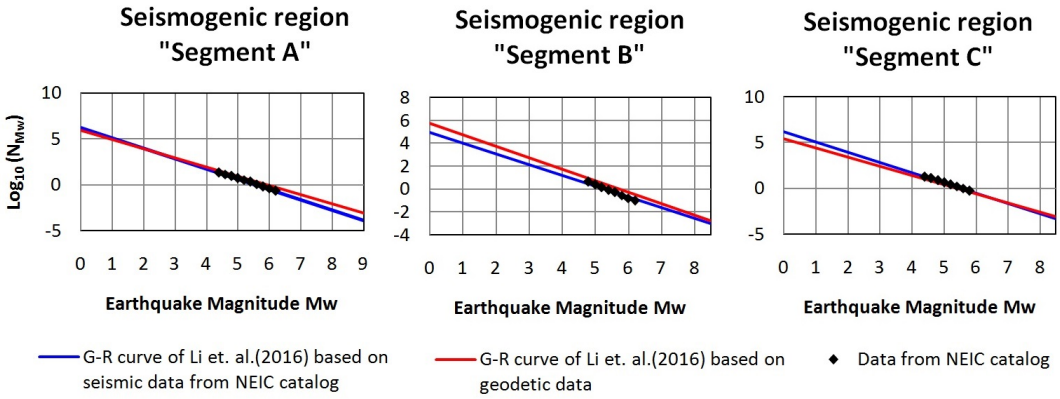


Figure 5.3: G-R curves of Li et al. (2016) based on seismic data from NEIC earthquake catalog (blue curve) and based on the geodetic data provided by Hsu et al. (2012) (red curve). Black diamonds are recurrences of earthquakes larger than different earthquake magnitudes, which are obtained from NEIC catalog.

computed by using the G-R curves and the following expression,

$$\lambda^{EQ}_{M_{Wj},x_i} = N(M_{Wj} - \tfrac{\Delta M_{Wj}}{2}) - N(M_{Wj} + \tfrac{\Delta M_{Wj}}{2}), \tag{5.3}$$

and the return period can be obtained by,

$$T^{EQ}_{M_{Wj},x_i} = \frac{1}{\lambda^{EQ}_{M_{Wj},x_i}}. \tag{5.4}$$

Note that different earthquake magnitudes are contained in each interval. Since the slip random field model generates earthquake samples with different magnitude, we adopt a ΔM_{Wj} which contains most of the random field sample magnitude variability. Fig. 5.4 shows the deviation from the expected earthquake magnitude of the samples, for different segments and magnitudes. According to these histograms, we have specified $\Delta M_{Wj} = 0.25$ units, in which most of the samples are contained for any expected earthquake magnitude. The minimum and maximum earthquake magnitude is specified differently for each of the three segments. The minimum earthquake magnitudes are chosen so the

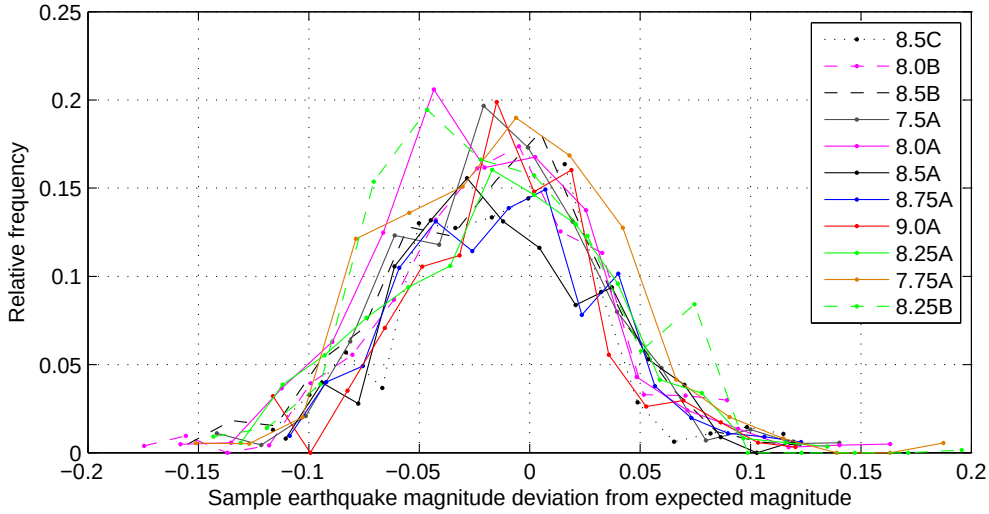


Figure 5.4: Histogram of the deviation of earthquake magnitudes from the expected magnitude as a result of the slip random field model. Different curves correspond to the histograms of different expected magnitudes and segments.

tsunamis generated by these earthquakes generate a significant impact at the assessed locations. We consider a minimum expected magnitude M_W of 7.5, 8.0 and 8.5 for segments A, B and C, respectively. As we mentioned before, the maximum earthquake magnitudes are specified according to the size of the seismogenic region. They are 9.0, 8.5 and 8.5 for segments A, B and C, respectively. The return periods $T^{EQ}_{M_{Wj},x_i}$ for different expected magnitudes and segments of this PTHA are shown in Table 5.2.

5.1.3 Earthquake sample generation.

Following the framework of Chapter 3 we generate synthetic earthquakes in which the slip and the centroid location are uncertain.

Expected Mw	Mw Range	Seismic based $T^{EQ}_{M_{Wj},x_i}$ [yr]	Geodetic based $T^{EQ}_{M_{Wj},x_i}$ [yr]
Segment A		a=6.2 ; b=1.12	a=5.96 ; b=1
7.5	7.375-7.625	242	59
7.75	7.625-7.875	460	106
8.0	7.875-8.125	877	188
8.25	8.125-8.375	1672	334
8.5	8.375-8.625	3185	594
8.75	8.625-8.875	6069	1056
9.0	8.875-9.125	11565	1879
Segment B		a=4.96 ; b=0.94	a=5.74 ; b=1
8.0	7.875-8.125	1139	312
8.25	8.125-8.375	1493	554
8.5	8.375-8.625	1956	986
Segment C		a=5.97 ; b=1.2	a=5.45 ; b=1
8.5	8.375-8.625	24103	1922

Table 5.2: Return periods $T^{EQ}_{M_{Wj},x_i}$ for the earthquakes considered in the PTHA, based on seismic data and geodetic data. The G-R parameters a and b are given for each segment and for each approach (seismic or geodetic based approaches).

The slip random field is built by specifying the rupture area using the scaling relations of Blaser et al. (2010) (see Fig. 2.2). We again adopt a Von Karman covariance function and a Log-normal marginal distribution. The correlation lengths and slip standard deviation are obtained from Raghukanth and Sangeetha (2016) scaling relations (see Fig. 2.10). Unlike the illustration case of Chapter 3, in this study case we specify a random centroid location along the strike direction for earthquake magnitudes smaller than the maximum. Hence, the centroid location in such cases is free to move in two directions. Table 5.3 shows the parameters used to build the earthquake samples, which were obtained from scaling relations in terms of the earthquake magnitude. The number of sub-faults was chosen to obtain a discretization error $\epsilon_d < 0.6\%$. The number

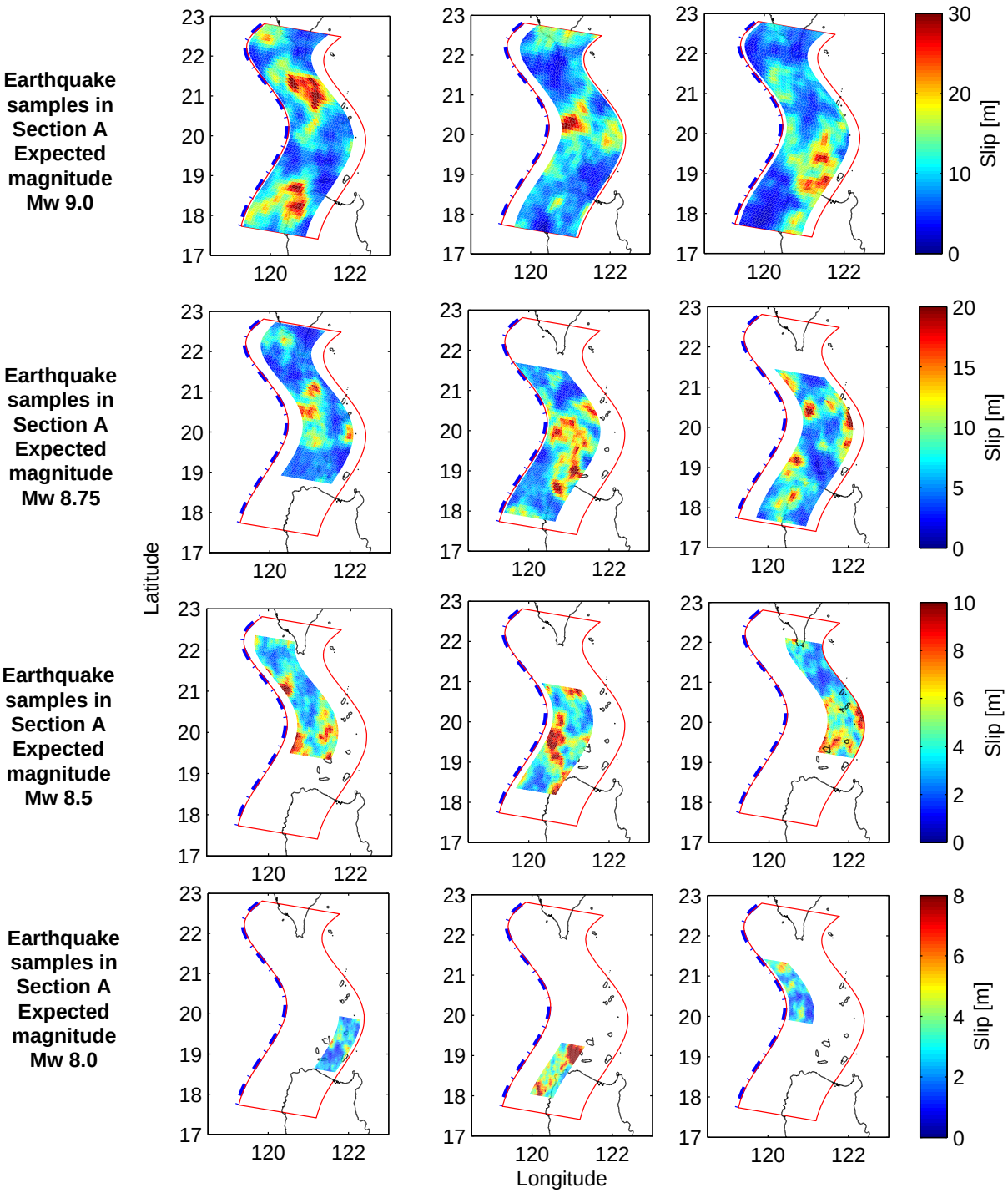
Expected M_w	Rupture Area L [km]	W[km]	Correlation a_x [km]	a_y[km]	σ_S [m]	$\bar{S}$ [m]	Sub-faults	σ_{M_W} []
Segment A (θ=10^o; max. depth=89 km; min. latitude=17.7^o)								
7.5	80	39	6.10	11.7	0.81	1.79	528	0.047
7.75	112	51	7.84	13.8	1.10	2.35	644	0.052
8.0	155	66	10.2	16.4	1.50	3.08	759	0.050
8.25	215	86	13.1	19.3	2.04	4.03	897	0.052
8.5	299	112	17.0	22.8	2.80	5.28	1058	0.044
8.75	414	146	22.0	27.0	3.81	6.93	1242	0.045
9.0	575	191	31.9	28.4	5.20	9.08	1449	0.042
Segment B (θ=0^o; max. depth=89 km; min. latitude=15.0^o)								
8.0	155	66	10.2	16.4	1.50	3.08	759	0.047
8.25	215	86	13.1	19.3	2.04	4.03	897	0.053
8.5	299	112	17.0	22.8	2.80	5.28	1058	0.051
Segment C (θ=330^o; max. depth=110 km; min. latitude=12.5^o)								
8.5	299	112	17.0	22.8	2.80	5.28	1058	0.044

Table 5.3: Parameters of the generated samples for each segment and magnitude. The last column informs about the earthquake magnitude standard deviation of samples, σ_{M_W}. An interval of $6\sigma_{M_W} \approx 0.25$ covers about the 99.7% of samples. This latter has been adopted as the magnitude interval $\Delta M_{W\,j}$ in the PTHA.

of K-L terms used for all the earthquake was 850. Three earthquake samples for each segment and some expected magnitudes are presented in Fig. 5.5. It is important to remark again that the generated samples are consistent with the probability properties originally defined for the slip random field.

5.2 Uncertainty propagation

One of the important goals of the methodology presented in this thesis is the uncertainty quantification of the tsunami response at the coast and in shallow water areas. The PTHA of Li et al. (2016) has two inconvenient simplifica-


Earthquake samples in Section A Expected magnitude Mw 9.0
Earthquake samples in Section A Expected magnitude Mw 8.75
Earthquake samples in Section A Expected magnitude Mw 8.5
Earthquake samples in Section A Expected magnitude Mw 8.0
Latitude
Longitude
Slip [m]
23
22
21
20
19
18
17
120
122
30
20
10
0
15
5
8
6
4
2

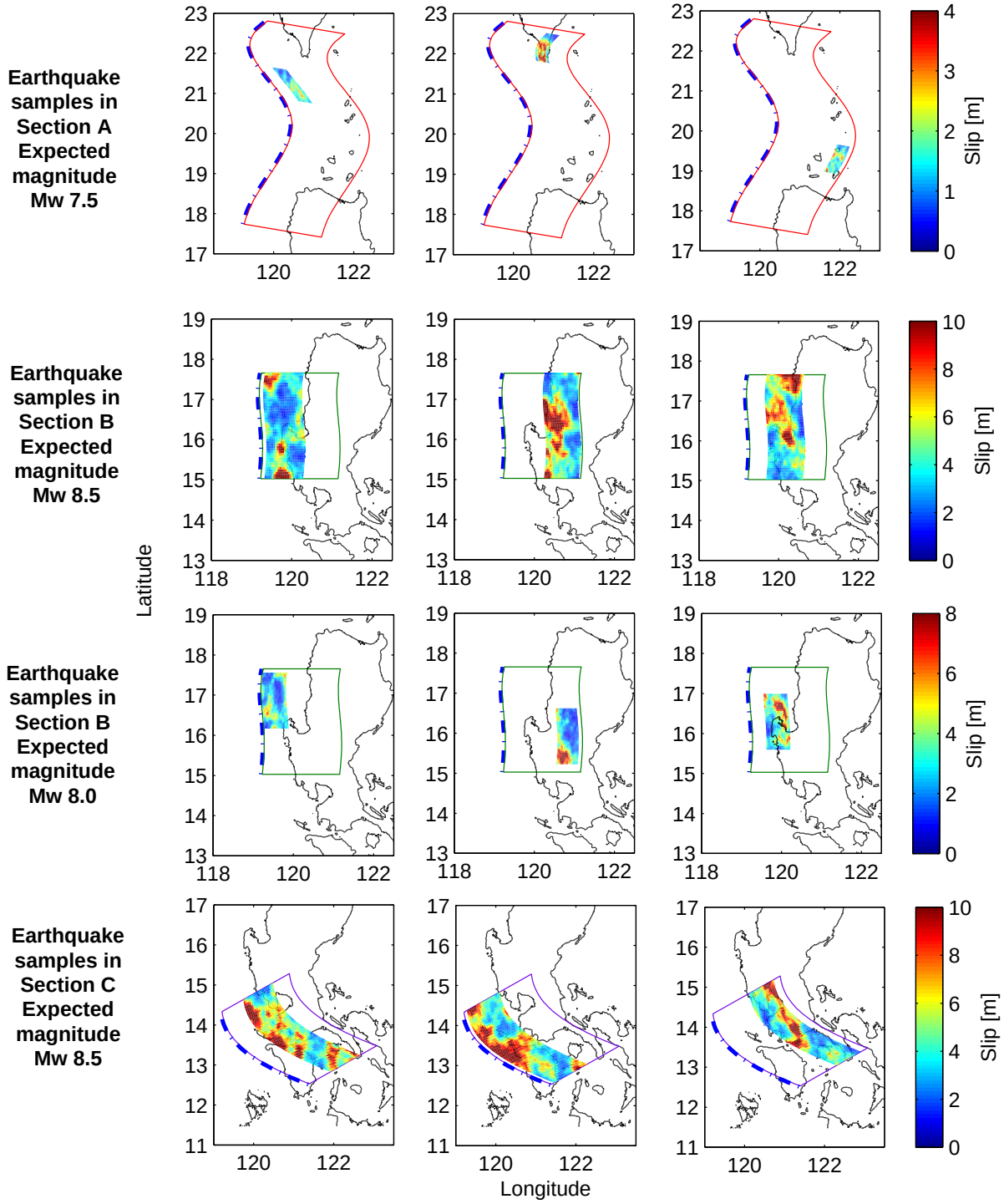


Figure 5.5: Three earthquake samples for each segment of the Manila Subduction Zone and some expected magnitude considered in the PTHA.

tions. First, their approach considers the linear shallow water wave model. This model is no longer valid in shallow areas since non linear inertia terms become comparable to the linear terms. Second, they consider a small number of earthquake scenarios and, therefore, might not capture all the earthquake variability in slip and location. Our methodology uses the non linear shallow water equations, which captures the tsunami response in shallow areas. The SROM model, on the other hand, allows to select an adequate set of earthquake samples which jointly capture the target slip and location statistical properties.

The SROM is implemented for the PTHA using the same procedure described in Chapter 3. We have generated 10,000 earthquake samples for each magnitude and seismogenic region (segment) defined above. The samples were then used to generate 10,000 vertical seafloor deformation, which were used to calculate the target probability properties of the seafloor deformation. Finally, n_{set}=100 sets of m=200 earthquakes were used to compute probabilities by means of the sub-optimal procedure. The set with the smallest discrepancies, with respect to the target probability properties of the seafloor displacement, was selected.

5.2.1 Tsunami model

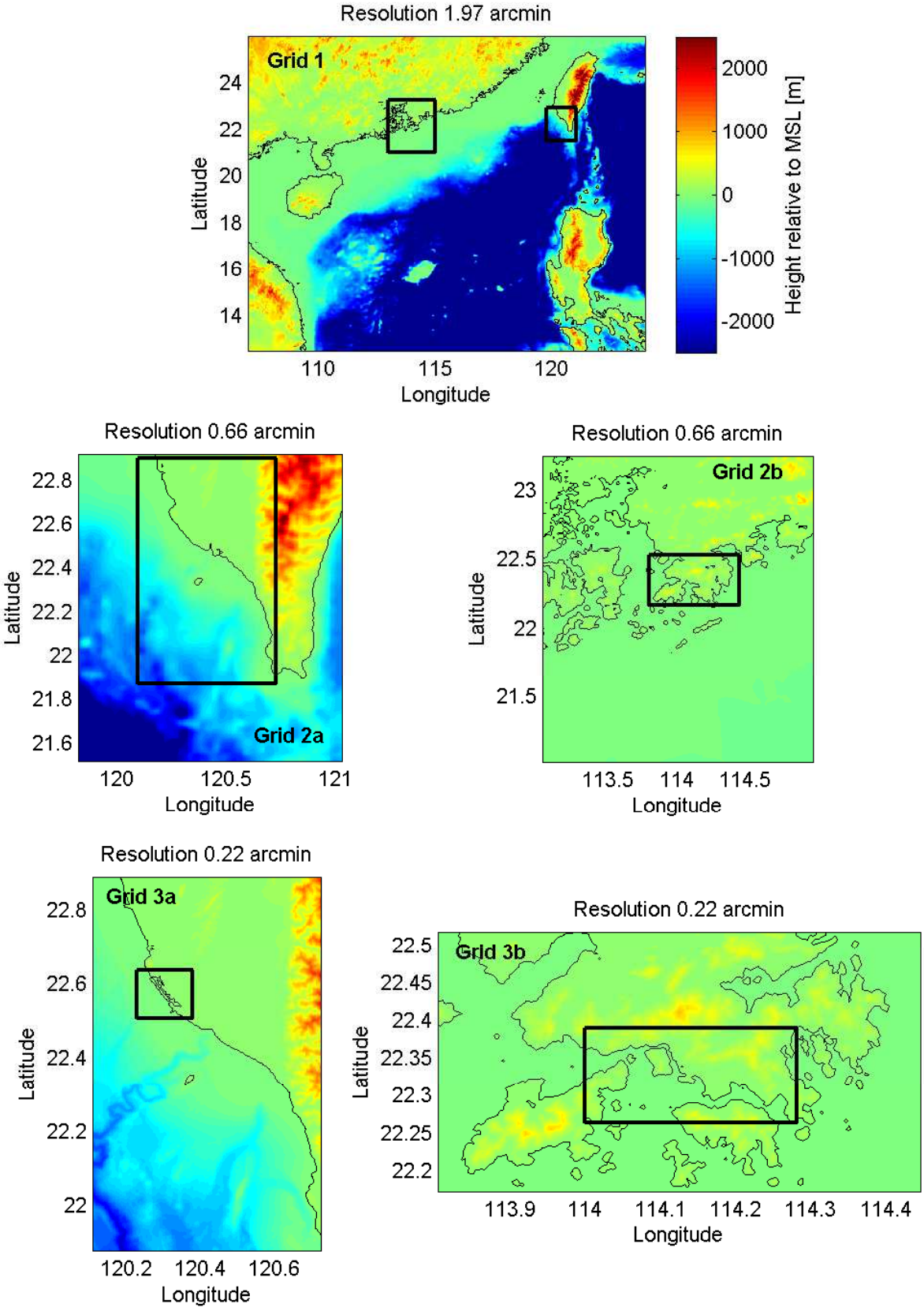
The set of m=200 vertical seafloor displacements are used as initial conditions for tsunami simulations in the SCS. Furthermore, we update the topo-bathymetry to account for the co-seismic deformation. As in the illustration case of Chapter 3, we use the tsunami model COMCOT. The model simulates bottom friction by adopting a quadratic friction law model with a constant Manning's number of

0.03. The model is set with nine grids which are used to capture the small scale response of the tsunami at the coast of Hong Kong and Kao Hsiung. For each area we use four nested grids over a coarse grid with 1.97 arc-min resolution. The finest grids in the assessed areas have a resolution of ~ 45 m. It is important to mention that the topo-bathymetry data used in this PTHA have equal or higher resolution than the numerical grids. The numerical grids are shown in Fig. 5.6. For grids 1, 2a and 2b we use GEBCO data of 0.5 arcmin resolution. For grid 3a we use a 200 m resolution DEM data of Taiwan. For grid 3b we use a combination of nautical charts of Hong Kong and SRTM topography data. For grids 4a and 5a we use a 40 m resolution DEM data of Taiwan, in combination to a nautical chart of ~ 30 m resolution in Kao Hsiung port (Hydrographic Office, 2016). For grid 4b and 5b we use the nautical charts of Hong Kong of ~ 20 m resolution and Lidar topography data with ~ 0.5 m resolution. It is important to remark that Kao Hsiung port is currently under a expansion plan which is modifying the topo-bathymetry at the south end of the port. This PTHA considers the current conditions of the Kao Hsiung port (i.e. 2016 Kao Hsiung Port nautical chart and DEM data with unknown date of elaboration).

5.2.2 Results

Maximum tsunami amplitude

We simulate tsunamis associated with the m=200 earthquake samples for each magnitude and seismogenic region segment presented in Table 5.2. Since co-seismic deformation might occur at the assessed locations, tsunamis will have a different impact whether the locations experience a subsidence or an uplift.


Resolution 1.97 arcmin
Grid 1
Latitude
Longitude
Height relative to MSL [m]
Resolution 0.66 arcmin
Grid 2a
Resolution 0.66 arcmin
Grid 2b
Resolution 0.22 arcmin
Grid 3a
Resolution 0.22 arcmin
Grid 3b

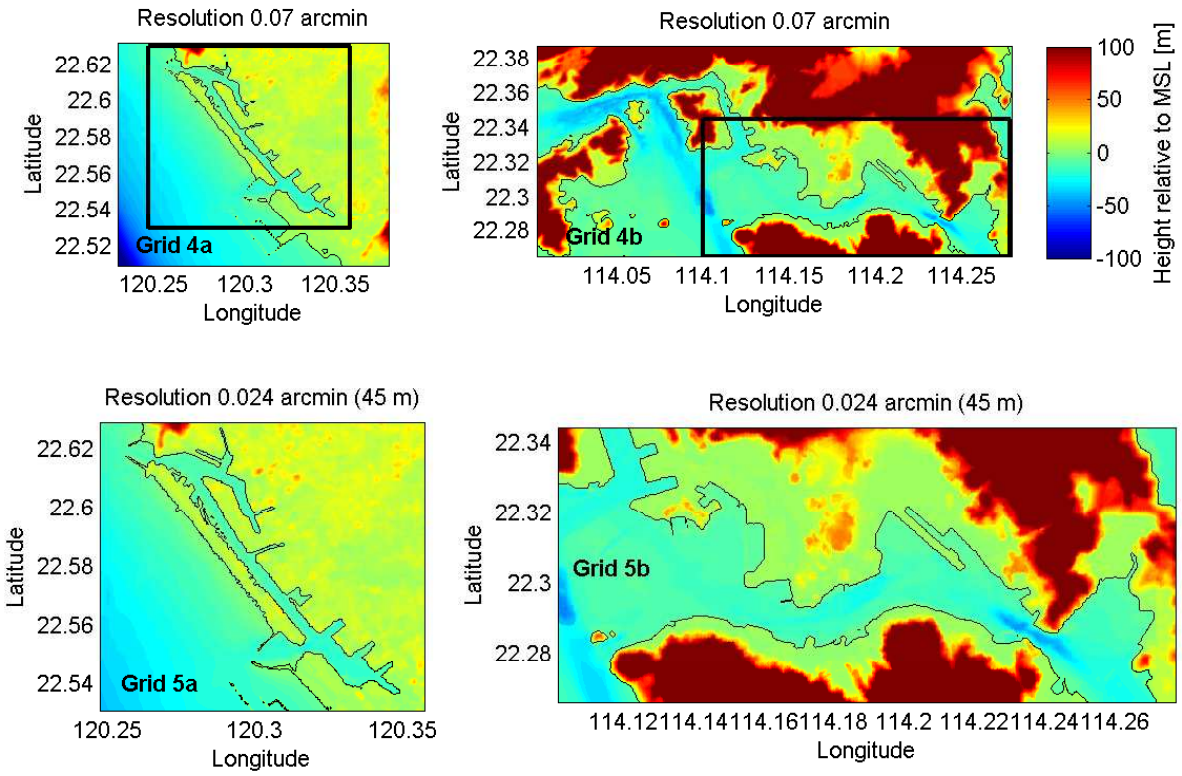


Figure 5.6: Grids (meshes) for the tsunami model in SCS. Heights refer to mean sea level (MSL). Black rectangles indicate the size of the next nested grid.

Thus, to quantify the overall effect of the tsunami at the coast, we correct the maximum tsunami amplitude values, by substracting the co-seismic vertical deformation at the assessed locations. By applying this correction, assessed locations which experiment subsidence (uplift) will perceive a larger (smaller) tsunami amplitude. The exceedance curves for each segment and magnitude are presented in Fig. 5.7 and Fig. 5.8.

The exceedance curves in Fig. 5.7 and Fig. 5.8 are combined with the earthquake return periods of Table 5.2 to compute tsunami return periods $T_R(h_{crit})$ of

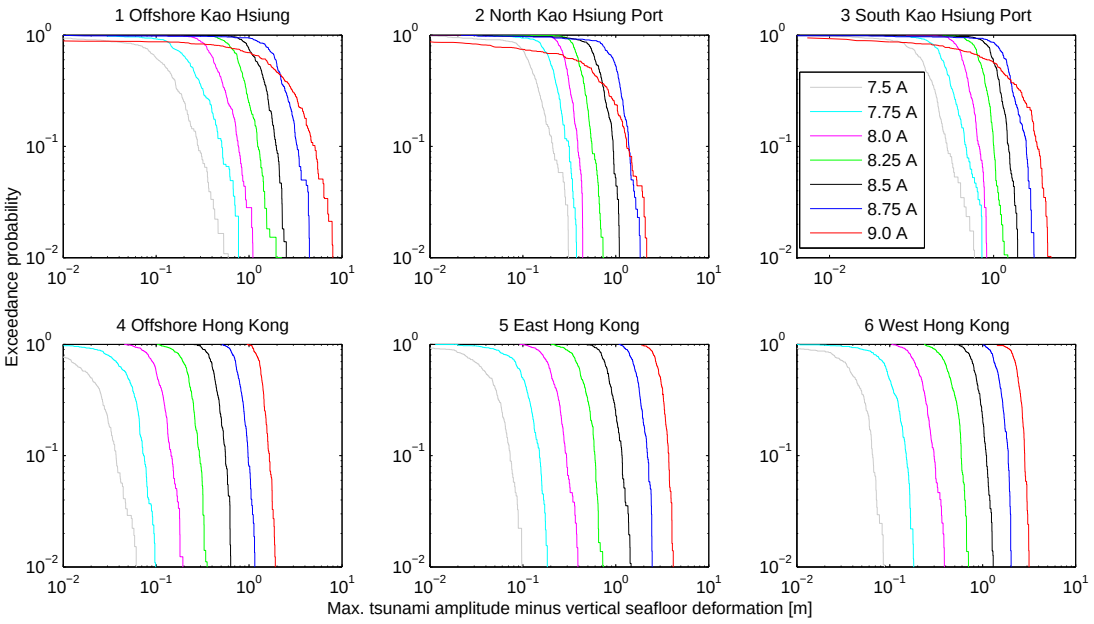


Figure 5.7: Exceedance curves of the maximum tsunami amplitude minus the vertical seafloor displacement at the assessed sites for different earthquake magnitudes in Segment A.

Eq. 2.28, where h_{crit} is defined as the maximum tsunami amplitude minus the co-seismic vertical deformation. The curves of h_{crit} in terms of $T_R(h_{crit})$ are called hazard curves herein. Fig. 5.9 shows the hazard curves for the six assessed locations in Kao Hsiung and Hong Kong. We show two curves for each location. The blue curves correspond to the hazard curves built with the earthquake return periods of the seismic based approach. The red curve, on the other hand, used earthquake return periods of the geodetic based approach. We observe that the earthquake recurrence models are associated with significant differences in the estimated return periods. In general, the PTHA with the geodetic based recurrence model presents the largest tsunami amplitudes for a given return period. We also observe low h_{crit} values and large return periods at the six assessed location. Station 1 in Kao Hsiung and station 5 in Hong Kong have the

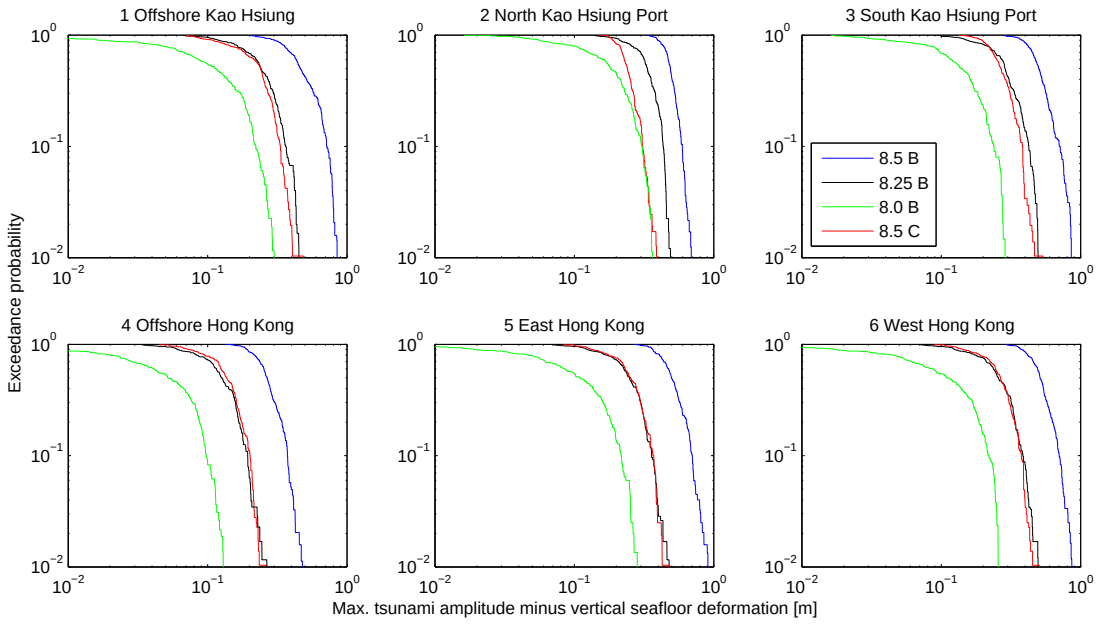


Figure 5.8: Exceedance curves of the maximum tsunami amplitude minus the vertical seafloor displacement at the assessed sites for different earthquake magnitudes in Segments B and C.

largest maximum tsunami amplitudes, which do not exceed 5 meters for return periods of 10,000 years.

Inundation

The inundation areas can be also assessed by following a probabilistic approach. Inundation occurs when the tsunami amplitude exceeds the topography height of previously dry areas. The topo-bathymetry height is defined as $-b$ in Fig. 2.4 and is referred to the still water level (SWL). By defining h_{crit} as the maximum tsunami amplitude for the assessment of inundation areas, we use Eq. 2.28 to compute $T_R(h_{crit} = -b)$ at each node of the grids 5a and 5b in Fig. 5.6. The solutions are interpreted as the return period of inundation by a tsunami,

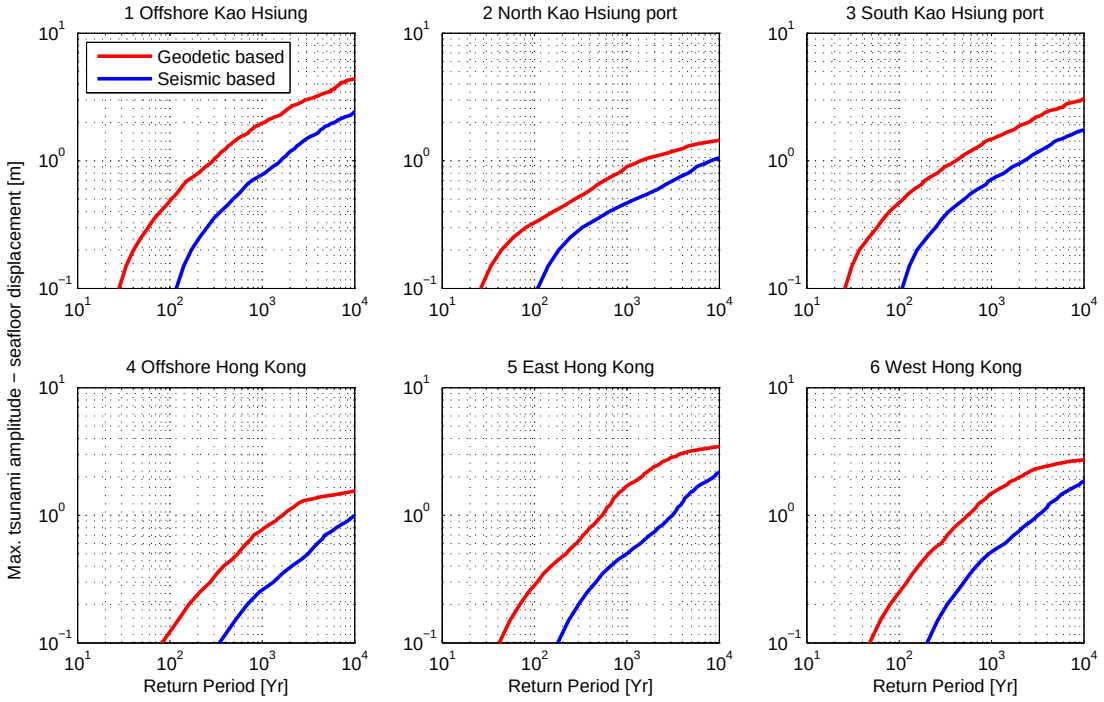


Figure 5.9: Hazard curves for the six assessed locations in Kao Hsiung and Hong Kong.

independent of the inundation depth.

The inundation maps are shown in Fig. 5.10 and the colors represent the logarithm of return periods. Note that the black zones, with return periods smaller than 1 year, correspond to the areas which are always flooded (i.e. areas with topo-bathymetry height below the still water level). Both ports are assessed in terms of the geodetic (left) and seismic (right) based earthquake recurrence models. We remark here that the inundation map calculation considers the co-seismic deformation of the assessed areas, which is significant in Kao Hsiung when earthquakes occur in segment A.

According to the geodetic and seismic based recurrence model approaches,

the port of Kao Hsiung has a significant inundation at the shore side of the port. The return periods associated with the geodetic based model are considerably smaller than those of the seismic based model. The south side of the port is significantly flooded with return periods of about 1,000 years, according to the geodetic based model. The north side of the port also present some inundation risk, but with return periods larger than 1,000 years, according to the geodetic based model. We observe a small inundation inside the harbor, with localized inundation areas at the south. The area of Hong Kong is almost unaffected by tsunamis generated in the Manila Subduction Zone. We only observe a small tsunami inundation hazard in the surroundings of Kai Tak cruise terminal (Lat. 22.31^o, Lon. 114.21^o) and some sections of the waterfront, with return periods of about 10,000 years according to the geodetic based model. The living area remains dry. It is important to mention that disparate tsunami hazard levels have been estimated in the literature for Kao Hsiung and Hong Kong. The literature shows many worst case scenario assessments (defined in chapter 1 and section 2.3) for the SCS. Most of these assessments (e.g. Megawati et al. (2009), Dao et al. (2009), Ren et al. (2015)) used a similar M_W 9.3 earthquake scenario in the Manila Subduction Zone, which was proposed by Megawati et al. (2009). Li et al. (2016), on the other hand, performed a PTHA in the SCS, which adopted the same earthquake recurrence models used this thesis. Their estimated tsunami hazard in Hong Kong and Kao Hsiung, however, are significantly larger than those estimated by us. The similarities of the earthquake scenarios employed in the mentioned studies suggest that the type of tsunami propagation model and its configuration (e.g. set up of nested grids and simulation of bottom friction) may cause the differences in the assessed tsunami hazard. In the next section we analyze the relevance of the tsunami propagation and configuration.

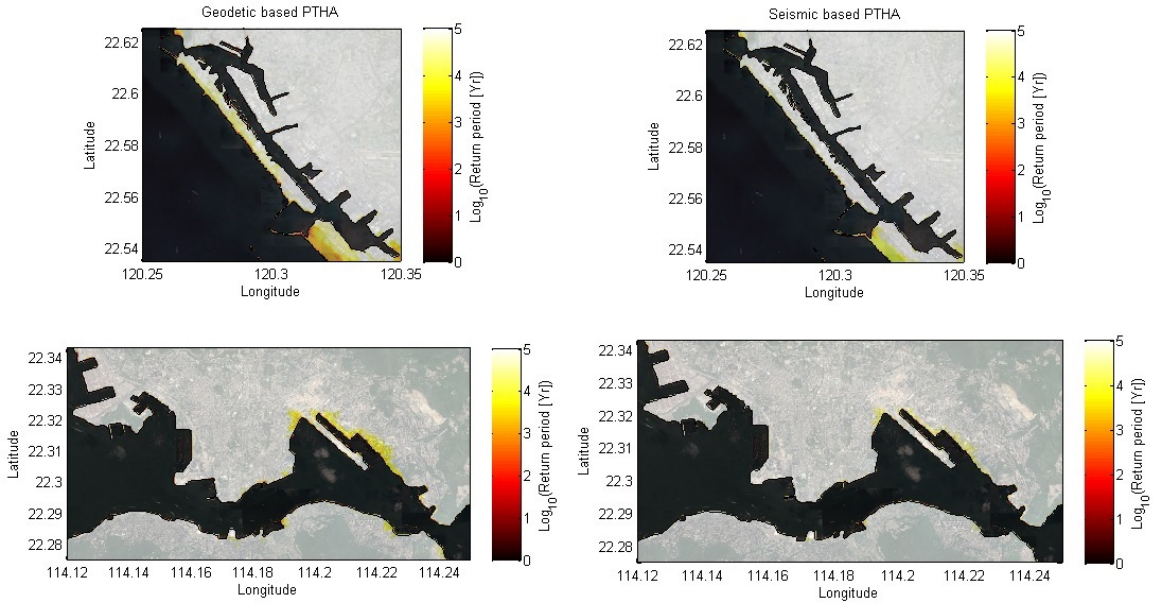


Figure 5.10: Inundation maps for Kao Hsiung (top panels) and Hong Kong (bottom panels). The left panels are result of the PTHA based on geodetic data and the right panels are results of the PTHA based on seismic data.

While the results in Kao Hsiung and Hong Kong show large return periods and thus a small tsunami hazard, some aspects have to be considered when the coastal hazard of these locations is assessed. First, the impact of tsunamis generated in the Manila Subduction Zone is very small and, therefore, other tsunami sources may become more relevant. For instance, Terry et al. (2017) suggest that landslide tsunami waves might be significantly large in coastal locations which are close to the landslide. Other marine hazards, such as tropical cyclones, might also be more relevant than the studied tsunamis. It is important to mention that Megawati et al. (2009) describe tsunami records in Hong Kong (in 1076 A.D.) and Taiwan (in 1781). The small hazard estimates of the present PTHA suggest that those past tsunami events were not generated by earthquakes in the Manila Subduction Zone, but by other sources. In fact, no

evidence of a large earthquake accompanies those tsunami records. Second, our PTHA does not consider water level variations due to tides. The tides ranges of the SCS (and open harbors) are in the order of 2 meters, which is comparable with the estimated tsunami amplitudes. The sea level rise (SLR) is not included, neither. However, we think that a the SLR is irrelevant for PTHA. The PTHA are commonly planned to be used in a relatively short window of time, e.g. the urban planning for the next 10 years or a project horizon of 30 years. The SLR in 10 or 30 years is much smaller than the tsunami amplitudes studied here. Finally, it is important to remark that the PTHA uses topo-bathymetry data which was surveyed recently. Anthropogenic modifications on the topo-bathymetry of ports are long term, but relevant. For instance, the expansion plan of the Kao Hsiung port (Fairplay, 2015) may significantly modify the inundation maps estimated in the south side of the port (Fig. 5.10). This latter aspect also demonstrates that PTHA are applicable to a short window of time (i.e. some decades) and as long as the current conditions remain unchanged.

5.3 Comparison with other tsunami hazard assessments in SCS.

As we mentioned above, different worst case scenario assessments have been conducted in the SCS using the earthquake scenario of Megawati et al. (2009), in which the three segments of Fig. 5.1 rupture at the same time. Megawati et al. (2009) used this earthquake scenario to estimate a maximum tsunami amplitude of $6 - 8$ m in Hong Kong. Dao et al. (2009) used the same earthquake scenario of Megawati et al. (2009) and a different tsunami propagation model. They,

however, estimated a maximum tsunami amplitude in Hong Kong of < 2m. Moreover, Ren et al. (2015) used a similar earthquake scenario and estimated a maximum tsunami amplitude of 3.7 m in Hong Kong. The fact that all the studies used similar (or the same) earthquake scenarios suggests that the differences are caused by the adoption of different tsunami propagation models and configurations. To assess the relevance of these latter, in section 5.3.1 we adopt the worst case earthquake scenario of Megawati et al. (2009) and perform a sensitivity analysis of the tsunami propagation model and configuration.

Li et al. (2016) conducted a PTHA for Hong Kong and Kao Hsiung (among other locations). A comparison with this study is relevant for our application case because they used same earthquake recurrence models but their results are significantly different. The relevant differences with both PTHA assessment are the following. (1) Both PTHA used different methodologies to generate earthquake samples. (2) We used several earthquake scenarios for each magnitude, while Li et al. (2016) only used a small number of samples, especially for large earthquakes (e.g. they use only 10 scenarios to represent M_W 9.0 earthquake in segment A). (3) We configured the model COMCOT with the non linear shallow water equations, bottom friction (with Manning 0.03) and nested grids, while Li et al. (2016) used linear shallow waters equations, no bottom friction and a single mesh with 1 arcmin resolution. We think the most relevant difference is related with the adoption of different tsunami propagation models and configurations. Thus, in section 5.3.2 we perform a PTHA using the same model configuration of Li et al. (2016) and we then compare it with the PTHA results described above.

5.3.1 Sensitivity of tsunami model and configuration.

We use the earthquake scenario of Megawati et al. (2009) to build a tsunami initial condition and to simulate eight sensitivity cases of the tsunami propagation model and configuration. The tsunami response is assessed in the locations presented in 5.1. Case 1 corresponds to the same model and configuration used in our PTHA of above. Cases 2 to 5 modify one model or configuration aspect, with respect to Case 1. The refinement of the coarsest mesh (case 2) has a negligible impact in tsunami responses. The use of a single coarse mesh (case 3), instead of the use of nested grids, has a significant effect in many assessed locations. Some locations experience a increasing of the maximum tsunami height in more than 50%. The removal of the bottom friction term (case 4) has also a significant effect. All assessed location experience an increase of the maximum tsunami height, varying between 0% and 70%. The adoption of a linear shallow water wave model (case 5), instead of a non linear shallow water wave model, has the greatest impact in the maximum tsunami amplitudes. It is important to mention that the bottom friction term is not included in the linear shallow water wave model. Some assessed locations experience an increase larger than 100%.

Case 6 corresponds to the simplest tsunami model and configuration. This case considers the linear shallow water wave model, a single mesh and no bottom friction. It is important to mention that this is the same model and configuration adopted by Li et al. (2016) in the tsunami simulations of their PTHA. Cases 7 to 9 correspond to sensitivity cases of case 6 in which we consider a coarser grid, nested grids and the non linear shallow water wave model. Note that cases 6, 7 and 8 have comparable maximum tsunami amplitudes, with some locations reaching more than 150% of variation with respect to Case 1. The case

adopting a non linear shallow water wave model (case 9), on the other hand, experience smaller maximum tsunami heights in the assessed locations, as compared with case 6.

This sensitivity analysis shows that the type of tsunami propagation model and configuration adopted have a relevant effect on the simulated tsunami responses. Comparisons between case 1 and case 5 and between case 6 and case 9 show that linear shallow water wave model simulates larger tsunami amplitudes, as compared with the non linear model. It is important to mention that the differences observed in this analysis are comparable to the differences between Megawati et al. (2009), Dao et al. (2009) and Ren et al. (2015).

5.3.2 PTHA using an alternative tsunami model configuration.

In this section we perform a second PTHA adopting the same tsunami model and configuration of Li et al. (2016), which is the linear shallow water wave equation, one grid of 1 arcmin resolution and no bottom friction. Since the grid resolution of this new PTHA is coarser, some of the assessed locations of the first PTHA (stations 1 to 6 in Table 5.1) are estimated to be on land. Thus, we assess tsunami amplitudes in six different locations (stations 7 to 12 in Table 5.1). The inundation, on the other hand, is not assessed for this second PTHA. The first PTHA, using nested grids, the non linear shallow water wave model and bottom friction, is called first PTHA, herein. The new PTHA, using the tsunami propagation model configuration of Li et al. (2016), is called second PTHA herein.

The segmented curves of Fig. 5.12 correspond to the results of h_{crit} (being

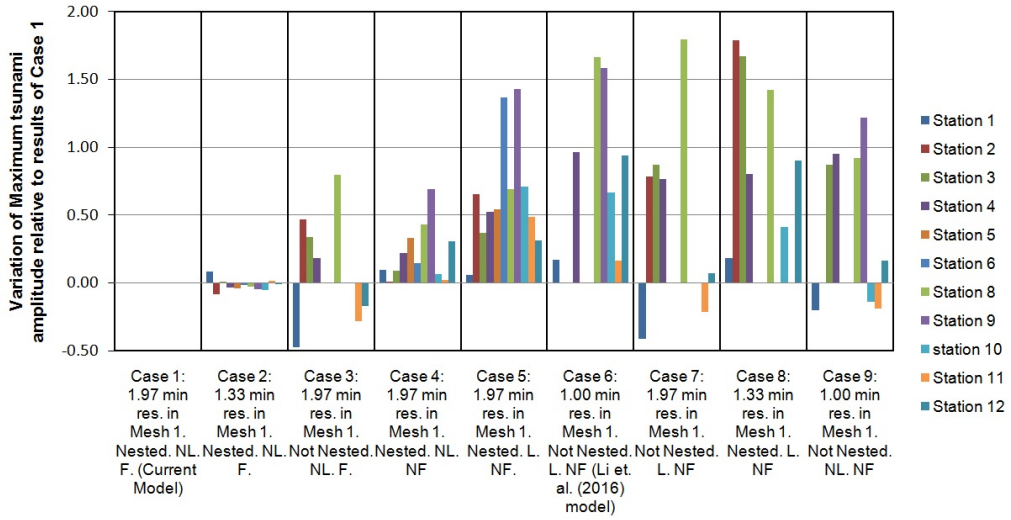


Figure 5.11: Comparison of the maximum tsunami heights at 11 assessed locations presented in Table 5.1 using 9 alternative configurations and tsunami models. The nomenclature is given as follows. **NF**: No friction, **F**:Friction with manning 0.03, **NL**: Non linear Shallow water equations, **L**: Linear shallow water equations, **Nested**: 8 nested curves to reach a resolution of ~ 40 m in assessed points with a coarsest grid of 1.33 arcmin or 1.97 arcmin. **Not nested**: a single mesh with resolution 1 arcmin or 1.97 arcmin. Station 7 was removed from the analysis because it is located on land according to the high resolution bathymetry data.

defined as the maximum tsunami amplitude minus the co-seismic vertical deformation) of the second PTHA in terms of $T_R(h_{crit})$. Fig. 5.13, on the other hand, shows the tsunami hazard curves presented by Li et al. (2016) using the seismic based earthquake recurrence model (green curves) and the geodetic based recurrence model (red curves). The hazard curves of the second PTHA at station 7 Hong Kong (segmented curves in 5.12) are comparable to those shown in the right panel of Fig. 5.13. Some differences can be seen in both studies, which are the result of the adoption of different earthquake sampling and uncertainty propagation methods. The hazard curves of the second PTHA at station 10 Kao

Hsiung (segmented curves in 5.12) show more differences with respect to the hazard curves of the left panel of Fig. 5.13. For instance, the maximum tsunami amplitudes for 10,000 years of return period differ in about three meters. We verified that this difference is caused by the seafloor displacement correction, which is not applied by Li et al. (2016). Relevant uplifts occur with the largest tsunami scenarios in Kao Hsiung, which reduce the values of Fig. 5.12.

To compare the first and second PTHA, we include the results of the first PTHA in Fig. 5.12 as solid curves. Since the assessed locations are different for both PTHA, we have plotted the results of the first PTHA at station 1 in Kao Hsiung and station 5 in Hong Kong. These correspond to the locations with largest tsunami amplitudes in the first PTHA. We observe significant discrepancies. The second PTHA estimates larger tsunami amplitudes. For instance, the maximum tsunami amplitude in Hong Kong for the first PTHA is ~1/2 of the maximum tsunami amplitude estimated with the second PTHA for return periods of 10,000 years.

As we also noted in the sensitivity analysis of section 5.3.1, the comparison of this section strongly suggests that the tsunami propagation model and its configuration plays an important role in the estimation of tsunami responses at the coast. In this particular case, the second PTHA with a simpler propagation model (i.e. single coarse mesh with linear shallow water wave equations) approximately doubles the estimates of the first PTHA in Hong Kong and Kao Hsiung. It is important to remark that the tsunami propagation model and configuration used in the first PTHA and the illustration cases of Chapter 3 and 4 were adopted based on some relevant arguments. First, wave amplitudes are comparable to water depths at the assessed locations and therefore, non linear

terms of the shallow water wave models are relevant. Second, the reduction of tsunami wavelenghts and their interaction with small length scales of the bathymetry require a refinement by means of nested grids. Third, high tsunami current speeds and shallow depths make bottom friction dissipation relevant (see Eq. 2.18) and should be included. Based on these arguments, the tsunami propagation model and configuration adopted in the first PTHA are expected to better simulate tsunami responses at the assessed locations, as compared with the second PTHA. It is important to remark, however, that the non linear shallow water wave model does not capture some relevant phenomena in shallow water areas, such as frequency dispersion and wave breaking. Thus, solutions of the non linear shallow water wave model are also subjected to errors.

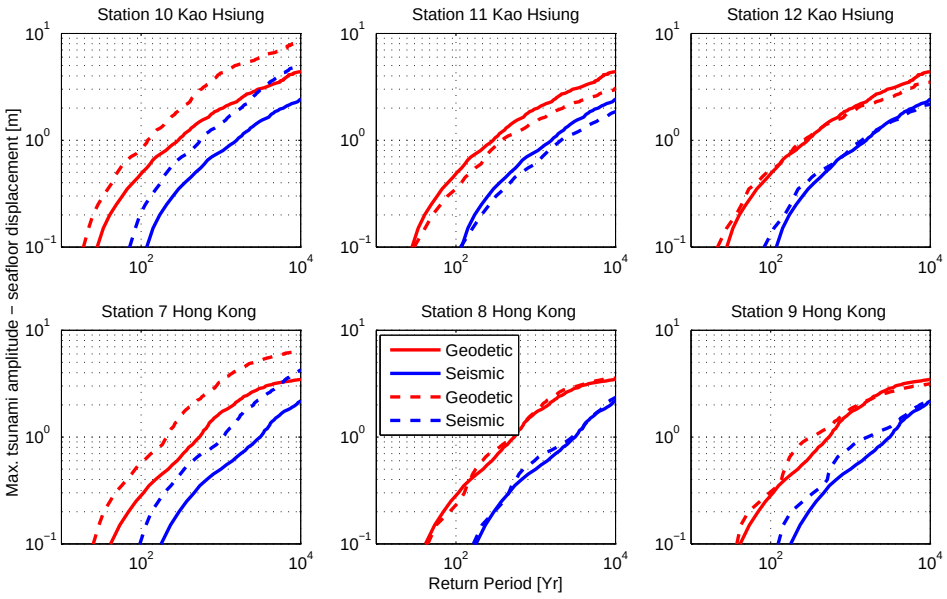


Figure 5.12: Segmented curves: Hazards curves of the second PTHA, using the same tsunami propagation model and configuration of Li et al. (2016). Solid curves: Hazard curves of the first PTHA, corresponding to the stations with largest tsunami amplitudes in Kao Hsiung (top panels) and Hong Kong (lower panels). The red curves are based on the geodetic approach for the earthquake recurrence model and the blue curves are based on the seismic approach.

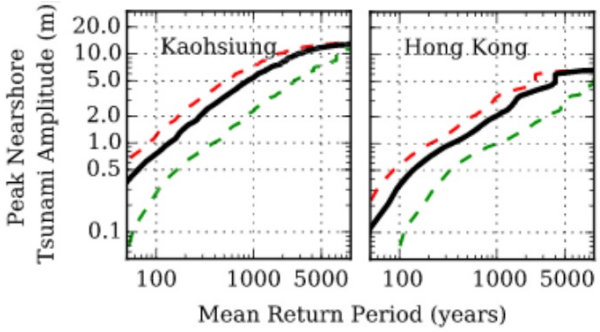


Figure 5.13: Hazard curves obtained by Li et al. (2016). The solutions are not corrected with the vertical seafloor displacement. The red segmented curves use the same earthquake recurrence model as the segmented red curve of Fig. 5.12. The green segmented curves use the same recurrence model as the blue segmented curves in Fig. 5.12. Black curves correspond to an average between the earthquake recurrence models.

CHAPTER 6

CONCLUSIONS

We describe the most relevant aspects of the proposed methodologies addressing the earthquake and bathymetry uncertainties. Then we describe the relevant results of the PTHA in South China Sea. Finally, we point out the future work required to improve the quantification of tsunami assessment uncertainties.

6.1 Relevant aspects of the methodology addressing uncertain earthquakes

Chapter 3 has proposed a methodology to characterize uncertainty in tsunami hazard assessments due to uncertain earthquake characteristics. The methodology assumes that the earthquake slip distribution and location have a random behavior, and they can be modeled as a random field and a random vector, respectively.

The uncertain slip distribution has been assumed to follow a Log-normal marginal distribution and a Von Karman covariance function (black curves in Fig. 3.17). We also assessed the uncertainty of tsunami responses by adopting an exponential marginal distribution for the slip (grey curves in Fig. 3.17), as no clear consensus was found in the literature about the best marginal distribution characterizing past earthquakes. The comparison between the black and grey curves in Fig. 3.17 demonstrates that the adoption of different target marginal distributions has a significant impact on tsunami responses. Thus, the slip marginal distribution to assess future tsunamis must be selected care-

fully and according to a better understanding of the statistical properties of past earthquakes. Different and more complex marginal distributions for the slip, such as those proposed by Gusev (2011) and Thingbaijam and Mai (2016), can also be adopted. However, the use of other marginal distributions require additional work to approximately determine covariance functions for the Gaussian random field. The scaling laws used to build the slip random field can also be revised and modified. We further remark that the slip distribution has been modeled as an homogeneous random field. Estimations of the characteristics of past earthquakes show that slip distribution exhibits a non-homogeneous behavior at the edges. Some studies, therefore, have adopted procedures tapering the slip at the rupture boundaries (e.g. LeVeque et al. (2016)). These procedures, however, adopt arbitrary taper functions and they lack of an understanding of the physics involved in such inhomogeneities. Hence, further investigation is needed to physically understand the spatial variability of the slip within the rupture area. The rupture location, on the other hand, has been assumed to follow a uniform distribution. From the literature review, though, it is noted that further investigation is required to determine more realistic probability properties.

The proposed sample generation method computes random samples in such a way that they are consistent with the target probability properties, defined for the slip distribution and location. The sample generation uses a K-L representation of Gaussian random fields, whose accuracy is controlled by two criteria defining the K-L truncation and the sub-fault discretization. These criteria allow to minimize the computational cost for a given accuracy. Another important feature is that the methodology admits non rectangular rupture areas, which are more appropriate for large earthquake events.

Using the Stochastic Reduced Order Model to propagate uncertainty, tsunami assessment uncertainties can be estimated with better accuracies than the classic Monte Carlo method with the same number of samples. Another advantage of the methodology is that the uncertainty propagation method is not intrusive and, therefore, requires only deterministic tsunami simulations.

It is important to mention that the proposed methodology only considers the slip distribution and location as uncertain earthquake parameters. The discrepancies between samples and tsunami records in Fig. 3.18, though, suggests that other sources of uncertainty are relevant in tsunami assessments. Note that methodology can be generalized to include further uncertain fault parameters or other inputs, if consistent probability properties of the uncertainty are provided.

6.2 Relevant aspects of the methodology addressing uncertain bathymetry

Chapter 4 has proposed a methodology to characterize the uncertainty of tsunami hazard assessments due to uncertain bathymetry. The uncertainty is associated with interpolation errors, which are assumed to be dominant. The methodology models the bathymetry as a random field, which is conditional to surveyed bathymetry data.

Unlike the earthquake characteristic uncertainties, the probability properties of bathymetry uncertainties have been addressed scarcely in the literature. Hence, this constitutes one the major challenges addressing bathymetry uncer-

tainties. In the illustration case of Chapter 4 we have adopted a Gaussian distribution with mean values assumed to be equal to a linear interpolation using surveyed bathymetry data. Based on the observed spectral properties of bathymetry transects (e.g. Mandelbrot (1967), Bell (1975)), we have assumed a Von Karman covariance function. The correlation lengths and variance, on the other hand, have been estimated qualitatively by comparing our model with high-quality bathymetry data at the site. It is important to mention, however, that tsunami assessment uncertainties are very sensitive to the variance and correlation lengths (see Fig. 4.11). Therefore, further investigation is required to validate consistent probability properties.

The methodologies for sample generation and uncertainty propagation follow a similar procedure as the methodology of Chapter 3. There are, however, some relevant peculiarities. The number of grid nodes in bathymetry meshes is typically $O(10^5 - 10^6)$. Consequently, the unconditional covariance matrix c_p of section 2.4.1 has $O(10^{10} - 10^{12})$ terms. This means that the computer RAM memory required to model bathymetry uncertainties exceeds 80 GB (considering 8 bytes per matrix term). The inversion of c_{p22} in Eq. 2.36 is equally demanding. Thus, a considerably large computational capacity is required to study real cases. Alternatively, numerical methods have to be sought to reduce the computational demand. The sample generation method, on the other hand, uses a K-L expansion. The number of K-L terms required to accurately generate bathymetry samples is $O(10^4)$ (an order larger than the number of K-L terms required in Chapter 3). As a consequence the computation of the eigenvalue decomposition of c_b and the solution of the SROM optimization problem require a significant computational capacity as well.

Results of the illustration case demonstrate that bathymetry uncertainties have a small impact on leading waves propagating through deep waters. Conversely, waves which propagate through shallow waters or reflect at the coast might be significantly affected by bathymetry uncertainties. The comparison with the 2014 Chilean tsunami, on the other hand, shows that neither earthquake characteristics nor bathymetry uncertainties can explain some phase discrepancies between tsunami simulations and measurements.

6.3 Relevant aspects of the PTHA in South China Sea

The PTHA application case of Chapter 5 considers uncertain earthquake characteristics. Bathymetry uncertainties were not considered because high quality data was provided in Kao Hsiung and Hong Kong. It is important to remark, though, that uncertain earthquake characteristics can be combined with uncertain bathymetry or any other uncertain input in SROM.

PTHA assessments with consideration of uncertain inputs require a set of cummulative density functions (called exceedance curves) of the tsunami response for different earthquake magnitudes and seismogenic regions. They also require information about earthquake recurrences. The first PTHA assessment of Chapter 5 adopts two earthquake recurrence models, which are proposed by Li et al. (2016). It is observed that PTHA results using these two recurrence models are significantly different. This observation demonstrates that recurrence models also contribute to the uncertainty of tsunami assessments. We also compare the first PTHA of Chapter 5 with the PTHA of Li et al. (2016) in SCS. Both studies use similar tsunami generation models and the same earthquake

recurrence models. The results in Hong Kong and Kao Hsiung, however, are significantly different. By mean of a sensitivity analysis of the model configuration we assessed the relevance of bathymetry mesh resolution, non linear terms of the shallow water wave equations and bottom friction terms. We demonstrate that those three aspects of the tsunami propagation model modify the results significantly and can explain the differences between the first PTHA of this thesis and the PTHA of Li et al. (2016).

Finally and according to our study, the maximum tsunami amplitudes in Hong Kong and Kao Hsiung are not significantly large and they are associated with long return periods. This means that other type of coastal hazards, such as those described in Terry et al. (2017), can be as relevant as earthquake tsunamis. Thus, the PTHA of Chapter 5 might not completely describe the inundation hazard of Hong Kong and Kao Hsiung.

6.4 Future work

Our proposed methodologies offer two relevant contributions to the existing methodologies addressing the tsunami hazard assessment with consideration of uncertain inputs. First, our modeled uncertain inputs are consistent with the specified target probability properties. Second, our uncertainty propagation method is more accurate than classic Monte Carlo simulations. Future work is needed to accurately describe the probability properties of the earthquake slip, rupture location and bathymetry.

The PTHA in Kao Hsiung and Hong Kong, with consideration of earthquakes in the Manila Subduction Zone, suggests that other coastal hazards

might be as relevant as the tsunamis studied in this thesis. Thus, further coastal hazards must be included in future hazard assessments. Another relevant aspect of the application case is that the tsunami propagation model and the earthquake recurrence model constitute important sources of uncertainty. Future PTHA with consideration of earthquake in the Manila Subduction Zone must analyze the relevance of these additional sources of uncertainty.

The comparison of the illustration cases with the measurements of the 2014 Chilean tsunami shows that discrepancies between tsunami simulations and measurements are not completely explained by the uncertainties of the earthquake slip, rupture location and bathymetry. Thus, other sources of uncertainty are relevant. One of these sources of uncertainties is associated with the tsunami model adopted in this thesis. The assumptions and simplifications adopted in the tsunami generation and propagation models might not hold under some circumstances and, therefore, they might lead to inaccurate simulations. For instance, section 2.1 described the anomalously large tsunami events known as "tsunami earthquakes". The hypothetical mechanisms explaining the generation of these tsunamis are not captured by the simplified tsunami generation model adopted in this thesis. Shallow water wave models are also inadequate to completely capture relevant phenomena of the tsunami propagation, such as wave frequency dispersion and wave breaking. These latter can be relevant during the inundation of coastal areas. Hence, not only input uncertainties but also tsunami models uncertainties have to be investigated in the future to completely quantify the uncertainty of tsunami hazard assessments.

BIBLIOGRAPHY

Ader, T., Avouac, J.-P., Liu-Zeng, J., Lyon-Caen, H., Bollinger, L., Galetzka, J., Genrich, J., Thomas, M., Chanard, K., Sapkota, S. N., et al. (2012). Convergence rate across the nepal himalaya and interseismic coupling on the main himalayan thrust: Implications for seismic hazard. *Journal of Geophysical Research: Solid Earth*, 117(B4).

Aki, K. (1966). Generation and propagation of G waves from the Niigata Earthquake of June 16, 1964: Part 1. A statistical analysis.

Allmann, B. P. and Shearer, P. M. (2009). Global variations of stress drop for moderate to large earthquakes. *Journal of Geophysical Research: Solid Earth*, 114(B1).

Ammon, C. J., Kanamori, H., and Lay, T. (2008). A great earthquake doublet and seismic stress transfer cycle in the central Kuril islands. *Nature*, 451(7178):561–565.

An, C. and Liu, P. L.-F. (2014). Characteristics of leading tsunami waves generated in three recent tsunami events. *Journal of Earthquake and Tsunami*, 8(03).

An, C., Sepúlveda, I., and Liu, P. L.-F. (2014). Tsunami source and its validation of the 2014 Iquique, Chile, earthquake. *Geophysical Research Letters*, 41(11):3988–3994.

Andrews, D. (1980). A stochastic fault model: 1. static case. *Journal of Geophysical Research: Solid Earth*, 85(B7):3867–3877.

Béjar-Pizarro, M., Socquet, A., Armijo, R., Carrizo, D., Genrich, J., and Simons, M. (2013). Andean structural control on interseismic coupling in the North Chile subduction zone. *Nature Geoscience*, 6(6):462–467.

Bell, T. (1975). Statistical features of sea-floor topography. In *Deep Sea Research and Oceanographic Abstracts*, volume 22, pages 883–892. Elsevier.

Betz, W., Papaioannou, I., and Straub, D. (2014). Numerical methods for the discretization of random fields by means of the Karhunen–Loève expansion. *Computer Methods in Applied Mechanics and Engineering*, 271:109–129.

Bird, P. and Kagan, Y. Y. (2004). Plate-tectonic analysis of shallow seismicity: Apparent boundary width, beta, corner magnitude, coupled lithosphere thickness, and coupling in seven tectonic settings. *Bulletin of the Seismological Society of America*, 94(6):2380–2399.

Blaser, L., Krüger, F., Ohrnberger, M., and Scherbaum, F. (2010). Scaling relations of earthquake source parameter estimates with special focus on subduction environment. *Bulletin of the Seismological Society of America*, 100(6):2914–2926.

Bletery, Q., Sladen, A., Delouis, B., and Mattéo, L. (2015). Quantification of tsunami bathymetry effect on finite fault slip inversion. *Pure and Applied Geophysics*, 172(12):3655–3670.

Bricker, J. D., Gibson, S., Takagi, H., and Imamura, F. (2015). On the need for larger Manning's roughness coefficients in depth-integrated tsunami inundation models. *Coastal Engineering Journal*, 57(02):1550005.

Brunn, S., Devriendt, L., Boulton, A., Derudder, B., and Witlox, F. (2016). Assessing the impacts of the global financial crisis on major and minor cities in South and Southeast Asia: A hyperlink analysis. In *Spatial Diversity and Dynamics in Resources and Urban Development*, pages 135–155. Springer.

Calder, B. and Mayer, L. A. (2003). Automatic processing of high-rate, high-

density multibeam echosounder data. *Geochemistry, Geophysics, Geosystems*, 4(6).

Carrier, G. and Greenspan, H. (1958). Water waves of finite amplitude on a sloping beach. *J. Fluid Mech*, 4(1):97–109.

Constantine, P. (2007). A primer on stochastic Galerkin methods. *Lecture Notes*.

Dao, M. H., Tkalich, P., Chan, E. S., and Megawati, K. (2009). Tsunami propagation scenarios in the South China Sea. *Journal of Asian Earth Sciences*, 36(1):67–73.

Davies, G., Horspool, N., and Miller, V. (2015). Tsunami inundation from heterogeneous earthquake slip distributions: Evaluation of synthetic source models. *Journal of Geophysical Research: Solid Earth*, 120(9):6431–6451.

De Risi, R. and Goda, K. (2016). Probabilistic earthquake–tsunami multi-hazard analysis: Application to the Tohoku Region, Japan. *Frontiers in Built Environment.*, 2:25.

Dorostian, A. and Zaré, M. (2009). An introduction to random field model to characterize complexity in near-field earthquake slip.

Engquist, B. and Majda, A. (1977). Absorbing boundary conditions for numerical simulation of waves. *Proceedings of the National Academy of Sciences*, 74(5):1765–1766.

Erdogan, S. (2009). A comparision of interpolation methods for producing digital elevation models at the field scale. *Earth surface processes and landforms*, 34(3):366–376.

Fairplay, I. (2015). Kaohsiung sees sluggish growth while expansion stays on track.

Field, R., Grigoriu, M., and Emery, J. (2015). On the efficacy of stochastic collocation, stochastic Galerkin, and stochastic reduced order models for solving stochastic problems. *Probabilistic Engineering Mechanics*, 41:60–72.

Galgana, G., Hamburger, M., McCaffrey, R., Corpuz, E., and Chen, Q. (2007). Analysis of crustal deformation in Luzon, Philippines using geodetic observations and earthquake focal mechanisms. *Tectonophysics*, 432(1):63–87.

Gallovič, F. and Brokešová, J. (2004a). The k-2 rupture model parametric study: example of the 1999 Athens earthquake. *Studia Geophysica et Geodaetica*, 48(3):589–613.

Gallovič, F. and Brokešová, J. (2004b). On strong ground motion synthesis with k- 2 slip distributions. *Journal of Seismology*, 8(2):211–224.

Geist, E. L. (2002). Complex earthquake rupture and local tsunamis. *Journal of Geophysical Research: Solid Earth*, 107(B5).

Geist, E. L. and Dmowska, R. (1999). Local tsunamis and distributed slip at the source. In *Seismogenic and Tsunamigenic Processes in Shallow Subduction Zones*, pages 485–512. Springer.

Geist, E. L. and Parsons, T. (2006). Probabilistic analysis of tsunami hazards. *Natural Hazards*, 37(3):277–314.

Goda, K. (2015). Effects of seabed surface rupture versus buried rupture on tsunami wave modeling: A case study for the 2011 Tohoku, Japan, Earthquake. *Bulletin of the Seismological Society of America*, 105(5):2563–2571.

Goda, K., Mai, P. M., Yasuda, T., and Mori, N. (2014). Sensitivity of tsunami wave profiles and inundation simulations to earthquake slip and fault geometry for the 2011 Tohoku earthquake. *Earth, Planets and Space*, 66(1):105.

Goda, K., Yasuda, T., Mori, N., and Mai, P. M. (2015). Variability of tsunami inundation footprints considering stochastic scenarios based on a single rupture model: Application to the 2011 Tohoku earthquake. *Journal of Geophysical Research: Oceans*, 120(6):4552–4575.

Goff, J. A. and Jordan, T. H. (1988). Stochastic modeling of seafloor morphology: Inversion of sea beam data for second-order statistics. *Journal of Geophysical Research: Solid Earth*, 93(B11):13589–13608.

González, F. I., Titov, V. V., Mofjeld, H. O., Venturato, A. J., Simmons, R. S., Hansen, R., Combellick, R., Eisner, R. K., Hoirup, D. F., Yanagi, B. S., et al. (2005). Progress in NTHMP hazard assessment. *Natural Hazards*, 35(1):89–110.

Grigoriu, M. (1995). *Applied non-gaussian processes: Examples, theory, simulation, linear random vibration, and MATLAB solutions(Book)*.

Grigoriu, M. (1998). Simulation of stationary non-Gaussian translation processes. *Journal of engineering mechanics*, 124(2):121–126.

Grigoriu, M. (2009). Reduced order models for random functions. application to stochastic problems. *Applied Mathematical Modelling*, 33(1):161–175.

Grigoriu, M. (2012). *Stochastic systems: uncertainty quantification and propagation*. Springer Science & Business Media.

Gusev, A. (2011). Statistics of the values of a normalized slip in the points of an earthquake fault. *Izvestiya, Physics of the Solid Earth*, 47(3):176–185.

Gusman, A. R., Murotani, S., Satake, K., Heidarzadeh, M., Gunawan, E., Watada, S., and Schurr, B. (2015). Fault slip distribution of the 2014 Iquique, Chile, earthquake estimated from ocean-wide tsunami waveforms and GPS data. *Geophysical Research Letters*, 42(4):1053–1060.

Hare, R., Eakins, B., and Amante, C. (2011). Modeling bathymetric uncertainty.

Hayes, D. E. and Lewis, S. D. (1984). A geophysical study of the Manila Trench, Luzon, Philippines: 1. crustal structure, gravity, and regional tectonic evolution. *Journal of Geophysical Research: Solid Earth*, 89(B11):9171–9195.

Hayes, G. (2011). Rapid source characterization of the 03-11-2011 Mw 9.0 off the Pacific Coast of Tohoku earthquake. *Earth Planets Space 63(7), 529534*.

Hayes, G. P., Bergman, E., Johnson, K. L., Benz, H. M., Brown, L., and Meltzer, A. S. (2013). Seismotectonic framework of the 2010 February 27 Mw 8.8 Maule, Chile earthquake sequence. *Geophysical Journal International*, 195(2):1034–1051.

Haynes, K. E., HSING, Y., and Stough, R. (1997). Regional port dynamics in the global economy: the case of Kaohsiung, Taiwan. *Maritime Policy & Management*.

Herrero, A. and Bernard, P. (1994). A kinematic self-similar rupture process for earthquakes. *Bulletin of the Seismological Society of America*, 84(4):1216–1228.

Hsu, Y.-J., Yu, S.-B., Song, T.-R. A., and Bacolcol, T. (2012). Plate coupling along the Manila subduction zone between Taiwan and northern Luzon. *Journal of Asian Earth Sciences*, 51:98–108.

Hydrographic Office, U. K. (2016). Nautical chart of Kao Hsiung.

IOC (1998). Post-tsunami survey field guide.

Iwasaki, S. (1982). Experimental study of a tsunami generated by a horizontal motion of a sloping bottom.

Jakobsson, M., Calder, B., and Mayer, L. (2002). On the effect of random errors in gridded bathymetric compilations. *Journal of Geophysical Research: Solid Earth*, 107(B12).

Kagan, Y. (1997). Seismic moment-frequency relation for shallow earthquakes: Regional comparison. *Journal of Geophysical Research: Solid Earth*, 102(B2):2835–2852.

Kajiura, K. (1963). The leading wave of a tsunami. *Bulletin of the Earthquake Research Institute*, 41:535–571.

Kanamori, H. (1972). Mechanism of tsunami earthquakes. *Physics of the earth and planetary interiors*, 6(5):346–359.

Kanamori, H. (1977). The energy release in great earthquakes. *Journal of geophysical research*, 82(20):2981–2987.

Kanamori, H. and Anderson, D. L. (1975). Theoretical basis of some empirical relations in seismology. *Bulletin of the Seismological Society of America*, 65(5):1073–1095.

Kanoglu, U. (2004). Nonlinear evolution and runup–rundown of long waves over a sloping beach. *Journal of Fluid Mechanics*, 513:363–372.

Kelker, D. (1970). Distribution theory of spherical distributions and a location-scale parameter generalization. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 419–430.

Kopp, H., Lange, D., Hannemann, K., Krabbenhoeft, A., Petersen, F., and Timmermann, A. (2016). RV SONNE Fahrtbericht/Cruise Report so244/2, GeoSEA: Geodetic Earthquake Observatory on the Seafloor, Antofagasta (Chile)–Antofagasta (Chile), 27.11.-13.12. 2015.

Kozdon, J. E. and Dunham, E. M. (2013). Rupture to the trench: Dynamic rupture simulations of the 11 March 2011 Tohoku earthquake. *Bulletin of the Seismological Society of America*, 103(2B):1275–1289.

Lay, T., Ammon, C. J., Kanamori, H., Rivera, L., Koper, K. D., and Hutko, A. R. (2010). The 2009 Samoa-Tonga great earthquake triggered doublet. *Nature*, 466(7309):964–968.

Lay, T., Yue, H., Brodsky, E. E., and An, C. (2014). The 1 April 2014 Iquique, Chile, Mw 8.1 earthquake rupture sequence. *Geophysical Research Letters*, 41(11):3818–3825.

LeVeque, R. J., Waagan, K., González, F. I., Rim, D., and Lin, G. (2016). Generating random earthquake events for probabilistic tsunami hazard assessment. *Pure and Applied Geophysics*, pages 1–22.

Li, L., Switzer, A. D., Chan, C.-H., Wang, Y., Weiss, R., and Qiu, Q. (2016). How heterogeneous coseismic slip affects regional probabilistic tsunami hazard assessment: A case study in the South China Sea. *Journal of Geophysical Research: Solid Earth*, 121(8):6250–6272.

Lin, J. (2015). Megathrust earthquake potential of the manila subduction system: revealed by the seismic moment tensor element Mrr 2.

Liu, P. L.-F., Cho, Y.-S., Briggs, M. J., Kanoglu, U., and Synolakis, C. E. (1995).

Runup of solitary waves on a circular island. *Journal of Fluid Mechanics*, 302:259–285.

Liu, P. L.-F., Wang, X., and Salisbury, A. J. (2009). Tsunami hazard and early warning system in South China Sea. *Journal of Asian Earth Sciences*, 36(1):2–12.

Liu, Y., Santos, A., Wang, S. M., Shi, Y., Liu, H., and Yuen, D. A. (2007). Tsunami hazards along Chinese coast from potential earthquakes in South China Sea. *Physics of the earth and planetary interiors*, 163(1):233–244.

Løvholt, F., Pedersen, G., Bazin, S., Kühn, D., Bredesen, R., and Harbitz, C. (2012). Stochastic analysis of tsunami runup due to heterogeneous coseismic slip and dispersion. *Journal of Geophysical Research: Oceans*, 117(C3).

Ludwig, W. J. (1970). The Manila Trench and west Luzon Trough–III. seismic-refraction measurements. In *Deep Sea Research and Oceanographic Abstracts*, volume 17, pages 553IN13563–562571. Elsevier.

Madsen, P. A. and Schaeffer, H. A. (2010). Analytical solutions for tsunami runup on a plane beach: single waves, n-waves and transient waves. *Journal of Fluid Mechanics*, 645:27–57.

Mai, P. M. and Beroza, G. C. (2000). Source scaling properties from finite-fault-rupture models. *Bulletin of the Seismological Society of America*, 90(3):604–615.

Mai, P. M. and Beroza, G. C. (2002). A spatial random field model to characterize complexity in earthquake slip. *Journal of Geophysical Research: Solid Earth*, 107(B11).

Mai, P. M. and Thingbaijam, K. (2014). SRCMOD: An online database of finite-fault rupture models. *Seismological Research Letters*, 85(6):1348–1357.

Mandelbrot, B. B. (1967). How long is the coast of Britain. *science*, 156(3775):636–638.

Megawati, K., Shaw, F., Sieh, K., Huang, Z., Wu, T.-R., Lin, Y., Tan, S. K., and Pan, T.-C. (2009). Tsunami hazard from the subduction megathrust of the South China Sea: Part I. source characterization and the resulting tsunami. *Journal of Asian Earth Sciences*, 36(1):13–20.

Mei, C. C., Stiassnie, M., and Yue, D. K.-P. (1989). *Theory and Applications of Ocean Surface Waves: Part 1: Linear Aspects Part 2: Nonlinear Aspects*. World Scientific.

Mei, C. C., Stiassnie, M., and Yue, D. K.-P. (2005). *Theory and applications of ocean surface waves: linear aspects*, volume 23. World scientific.

Meinig, C., Stalin, S. E., Nakamura, A. I., and Milburn, H. B. (2005). Real-time deep-ocean tsunami measuring, monitoring, and reporting system: The NOAA DART II: description and disclosure.

Métois, M., Socquet, A., Vigny, C., Carrizo, D., Peyrat, S., Delorme, A., Maureira, E., Valderas-Bermejo, M.-C., and Ortega, I. (2013). Revisiting the North Chile seismic gap segmentation using GPS-derived interseismic coupling. *Geophysical Journal International*, 194(3):1283–1294.

Mofjeld, H., Titov, V., González, F., and Newman, J. (2001). Tsunami scattering provinces in the Pacific Ocean. *Geophysical research letters*, 28(2):335–337.

Molnar, P. (1979). Earthquake recurrence intervals and plate tectonics. *Bulletin of the Seismological Society of America*, 69(1):115–133.

Okada, Y. (1985). Surface deformation due to shear and tensile faults in a half-space. *Bulletin of the seismological society of America*, 75(4):1135–1154.

Okal, E. A., Synolakis, C. E., and Kalligeris, N. (2011). Tsunami simulations for regional sources in the South China and adjoining seas. *Pure and applied geophysics*, 168(6-7):1153–1173.

Raghukanth, S. and Sangeetha, S. (2016). A stochastic model for earthquake slip distribution of large events. *Geomatics, Natural Hazards and Risk*, 7(2):493–521.

Ren, Z.-Y., Zhao, X., and Liu, H. (2015). Dispersion effects on tsunami propagation in South China Sea. *Journal of Earthquake and Tsunami*, 9(05):1540001.

Rhie, J., Dreger, D., Bürgmann, R., and Romanowicz, B. (2007). Slip of the 2004 Sumatra–Andaman earthquake from joint inversion of long-period global seismic waveforms and GPS static offsets. *Bulletin of the Seismological Society of America*, 97(1A):S115–S127.

Ross, S. M. (2014). *Introduction to probability models*. Academic press.

Roux, M. C. R. (2016). *Hong Kong High DensCity: merging urban scale with human scale*. PhD thesis, Carleton University Ottawa.

Satake, K. (1994). Mechanism of the 1992 Nicaragua tsunami earthquake. *Geophysical Research Letters*, 21(23):2519–2522.

Satake, K., Fujii, Y., Harada, T., and Namegaya, Y. (2013). Time and space distribution of coseismic slip of the 2011 Tohoku earthquake as inferred from tsunami waveform data. *Bulletin of the seismological society of America*, 103(2B):1473–1492.

Scholz, C. H. and Campos, J. (2012). The seismic coupling of subduction zones revisited. *Journal of Geophysical Research: Solid Earth*, 117(B5).

Sepúlveda, I. and Liu, P. L.-F. (2016). Estimating tsunami runup with fault plane parameters. *Coastal Engineering*, 112:57–68.

Shearer, P. M. (2009). *Introduction to seismology*. Cambridge University Press.

SHOA (2009). Atlas hidrografico de la Armada de Chile, 7th edition.

Shuto, N., Imamura, F., Yalciner, A., and Ozyurt, G. (2006). TUNAMI N2: Tsunami modeling manual. *Iwate Prefectural University*.

Sipkin, S., Person, W., and Presgrave, B. (2000). Earthquake bulletins and catalogs at the USGS National Earthquake Information Center. *IRIS Newsletter*, 2000(1):2–4.

Smith, W. and Wessel, P. (1990). Gridding with continuous curvature splines in tension. *Geophysics*, 55(3):293–305.

Song, Y. T., Mohtat, A., and Yim, S. C. (2017). New insights on tsunami genesis and energy source. *Journal of Geophysical Research: Oceans*.

Strasser, F. O., Arango, M., and Bommer, J. J. (2010). Scaling of the source dimensions of interface and intraslab subduction-zone earthquakes with moment magnitude. *Seismological Research Letters*, 81(6):941–950.

Synolakis, C. E. (1987). The runup of solitary waves. *Journal of Fluid Mechanics*, 185:523–545.

Tanioka, Y. and Satake, K. (1996). Tsunami generation by horizontal displacement of ocean bottom. *Geophysical Research Letters*, 23(8):861–864.

Tanioka, Y. and Seno, T. (2001). Sediment effect on tsunami generation of the 1896 Sanriku tsunami earthquake. *Geophysical Research Letters*, 28(17):3389–3392.

Tassara, A. and Echaurren, A. (2012). Anatomy of the Andean subduction zone: three-dimensional density model upgraded and compared against global-scale models. *Geophysical Journal International*, 189(1):161–168.

Te Chow, V. (1959). *Open channel hydraulics*. McGraw-Hill Book Company, Inc; New York.

Terry, J. P., Winspear, N., Goff, J., and Tan, P. H. (2017). Past and potential tsunami sources in the South China Sea: A brief synthesis. *Earth-Science Reviews*.

Thingbaijam, K. K. and Mai, P. M. (2016). Evidence for truncated exponential probability distribution of earthquake slip. *Bulletin of the Seismological Society of America*.

Thio, H. K., Somerville, P., and Ichinose, G. (2007). Probabilistic analysis of strong ground motion and tsunami hazards in Southeast Asia. *Journal of Earthquake and Tsunami*, 1(02):119–137.

Tinti, S. and Tonini, R. (2005). Analytical evolution of tsunamis induced by near-shore earthquakes on a constant-slope ocean. *Journal of Fluid Mechanics*, 535:33–64.

Titov, V. V. and Gonzalez, F. (1997). *Implementation and testing of the method of splitting tsunami (MOST) model*. US Department of Commerce, National Oceanic and Atmospheric Administration, Environmental Research Laboratories, Pacific Marine Environmental Laboratory.

Todorovska, M. I. and Trifunac, M. D. (2001). Generation of tsunamis by a slowly spreading uplift of the sea floor. *Soil Dynamics and Earthquake Engineering*, 21(2):151–167.

Tselentis, G.-A., Gkika, F., and Sokos, E. (2006). Tsunami hazards associated with the Perachora fault at eastern Corinth Gulf, Greece. *Bulletin of the Seismological Society of America*, 96(5):1649–1661.

Wang, X. (2009). User manual for COMCOT version 1.7 (first draft). *Cornell University*.

Weatherall, P., Marks, K., Jakobsson, M., Schmitt, T., Tani, S., Arndt, J. E., Rovere, M., Chayes, D., Ferrini, V., and Wigley, R. (2015). A new digital bathymetric model of the world's oceans. *Earth and Space Science*, 2(8):331–345.

Weeks, K., Mandal, P., and Sen, K. (2017). Advancements in technology and potential impacts on port automations decisions: The case of port of singapore. In *Entrepreneurship in Technology for ASEAN*, pages 127–137. Springer.

Wells, D. L. and Coppersmith, K. J. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the seismological Society of America*, 84(4):974–1002.

Williams, C. N., Cornford, S. L., Jordan, T. M., Dowdeswell, J. A., Siegert, M. J., Clark, C. D., Swift, D. A., Sole, A., Fenty, I., and Bamber, J. L. (2017). Generating synthetic fjord bathymetry for coastal greenland.

Winckler, P., Reyes, M., and Sepúlveda, I. (2010). Observaciones de campo del tsunami del 27 de febrero de 2010 en Isla Robinson Crusoe, Archipiélago Juan Fernndez. *Revista de la Sociedad Chilena de Ingeniería Hidráulica*, 5(1):1–12. ISSN 0716-3746.

Yagi, Y., Okuwaki, R., Enescu, B., Hirano, S., Yamagami, Y., Endo, S., and Komoro, T. (2014). Rupture process of the 2014 Iquique Chile earthquake in

relation with the foreshock activity. *Geophysical Research Letters*, 41(12):4201–4206.

Yamazaki, Y., Cheung, K. F., Kowalik, Z., Lay, T., and Pawlak, G. (2012). Neowave. In *Proceedings and Results of the 2011 NTHMP Model Benchmarking Workshop, NOAA, Galveston, Texas*.

Yap, W. Y., Lam, J. S., and Notteboom, T. (2006). Developments in container port competition in East Asia. *Transport Reviews*, 26(2):167–188.

Zambo, S. J., Elmore, P. A., Perkins, A. L., and Bourgeois, B. S. (2015). Uncertainty estimation for sparse data gridding algorithms. In *Proceedings of the US Hydro Conference, National Harbor, MD, USA*, pages 16–19.